Centers for Medicare & Medicaid Services

R & D Projects Database

CMS CENTERS for MEDICARE & MEDICAID SERVICES

PDF TRACKING FORM

Tracking Job Number	*63*	Printed:	04/23/2002
Document Title	Per-Episode Prospective Payment for Medicare Home Health Care Sharply Reduces Service Use		
Author	Christopher Trenholm	Document Date	30-Dec-00
Document Description:	Interim Report		

Project Specifics

SR Number	94-082
Project Number	500-94-0062
Project Title	Evaluation of Phase II of the Home Health Agency Prospective Payment Demonstration
Awardee	Mathematica Policy Research
Project Officer	Meadow, Ann
Principle Investigator	Cheh, Valerie
Funding Level	4013874
Start Date	09/30/1994
End Date	04/30/2001
Project Description:	This contract is evaluating Phase II of the Home Health Agency (HHA) Prospective Payment Demonstration, under which HHAs are paid on a prospective basis for an episode of care reimbursed by the Medicare program. (Phase I tested per-visit prospective payme

Initiated (CMS):	sgalloway	Date:	23-Apr-02	Accepted (IQ):		Date:	
Released (IQ):		Date:		Received (Sub):		Date:	
Finished (Sub):		Date:		QC'd (IQ):		Date:	
Returned (IQ):		Date:		pproved (CMS):		Date:	

Centers for Medicare & Medicaid Services

R & D Projects Database

PDF TRACKING FORM

Tracking Job Number ***121*** Printed: 5/15/02

Document Title: Per-Episode Prospective Payment for Medicare Home Health Care Sharply Reduces Service Use

Author: Christopher Trenholm

Document Date: 30-Dec-00

Document Description: Interim Report #6

Project Specifics

SR Number: 94-082

Project Number: 500-94-0062

Project Title: Evaluation of Phase II of the Home Health Agency Prospective Payment Demonstration

Awardee: Mathematica Policy Research

Project Officer: Meadow, Ann

Principle Investigator: Cheh, Valerie

Funding Level: 4013874

Start Date: 9/30/94

End Date: 4/30/01

Project Description: This contract is evaluating Phase II of the Home Health Agency (HHA) Prospective Payment Demonstration, under which HHAs are paid on a prospective basis for an episode of care reimbursed by the Medicare program. (Phase I tested per-visit prospective payme

Initiated (CMS):	sgalloway	Date:	15-May-02	→	Accepted (IQ):		Date:	
Released (IQ):		Date:		→	Received (Sub):		Date:	
Finished (Sub):		Date:		→	QC'd (IQ):		Date:	
Returned (IQ):		Date:		→	pproved (CMS):		Date:	

94_082_01_TJN63_Per...

Contract No.: 500-94-0062
MPR Reference No.: 8246-121

Per-Episode Prospective Payment for Medicare Home Health Care Sharply Reduces Service Use

December 2000

Christopher Trenholm

Submitted to:

Health Care Financing Administration
7500 Security Blvd, (C-3-21-06)
Baltimore, MD 21244-1850

Project Officer:
Ann Meadow

Submitted by:

Mathematica Policy Research, Inc.
P.O. Box 2393
Princeton, NJ 08543-2393
(609) 799-3535

Project Director:
Valerie Cheh

Principal Investigator:
Randall Brown

ACKNOWLEDGMENTS

In preparing this report, I received support from many colleagues for which I am grateful: Valerie Cheh and Randall Brown contributed their policy expertise and offered extensive and thoughtful comments throughout the research; Judith Wooldridge provided an excellent critical review; Barbara Phillips contributed to the background chapters through previous work; Elizabeth Stuart and Amy Zambrowski provided comprehensive and outstanding programming support; Marjorie Mitchell adeptly produced the report with assistance from Jennifer Baskwell, Denise Dunn, Cindy McClure, Jill Miller, and Jane Nelson; and Laura Berenson provided skilled editorial support with help from Walter Brower and Patricia Ciaccio. Ann Meadow of HCFA provided very useful comments to an earlier draft of the report. Any errors or omissions in this report are my own.

CONTENTS

Chapter **Page**

EXECUTIVE SUMMARY xv

I THE PER-EPISODE HOME HEALTH DEMONSTRATION AND EVALUATION 1

A. THE MEDICARE HOME HEALTH BENEFIT 2

B. THE PER-EPISODE DEMONSTRATION 4

1. Demonstration Payment and Incentives 6
2. Other Demonstration Procedures 9

C. GUIDE TO THE REST OF THIS REPORT 12

II DATA 13

A. CONSTRUCTING THE MAIN DATA FILE 14

1. Identifying Individual Patient Episodes (Records) 15
2. Creating Outcome Variables 16
3. Creating Control Variables 16
4. Additional Control Variables 22

B. FORMING THE ANALYSIS SAMPLE 24

1. Excluding Agencies 24
2. Excluding Individual Admissions 26
3. Analysis Samples 27

C. SUMMARY STATISTICS FOR THE MAIN CONTROL VARIABLES 31

1. Characteristics of the Sample 31
2. Implications for the Analysis of Program Impacts 36

CONTENTS *(continued)*

Chapter **Page**

III METHODS 39

A. UNITS OF ANALYSIS 39

1. Problem with Investigating Impacts Within the Patient Episode 39
2. Main Unit of Analysis: The "Patient Year" 42

B. STATISTICAL TECHNIQUES 46

1. Regression Models for Estimating Overall Impacts 47
2. Estimating Subgroup Impacts 49
3. Investigating Differences in Impacts Between Demonstration Years 52
4. Hypothesis Tests for the Impact Estimates 53
5. Design Effects 54
6. Weighting 55
7. Robustness Checks 56

IV IMPACT OF PROSPECTIVE PAYMENT ON THE USE OF SERVICES 63

A. EXPECTED EFFECTS 63

B. IMPACTS ON SERVICE USE IN THE YEAR AFTER ADMISSION . . . 66

C. IMPACTS ON SPECIFIC SERVICES DURING THE PATIENT-YEAR 73

1. Impacts During the At-Risk Period (First 120 Days) 73
2. Impacts After the At-Risk Period (120 Days) 77
3. Summary 80

D. IMPACTS ON THE DURATION AND TIMING OF SERVICES 81

E. ANALYSIS OF THE PATIENT EPISODE 86

1. Impacts on Key Episode Characteristics 87
2. Description of the Patient Episode Under Prospective Payment 91
3. Summary 98

CONTENTS *(continued)*

Chapter **Page**

IV *(continued)*

F. ROBUSTNESS TESTS 100

  1. Impacts Without Sample Weights 101
  2. Controlling for Detailed Patient Characteristics 103
  3. Effects of Outliers 105
  4. Effects of Censoring 107
  5. Influence of Unobserved Variables 108
  6. Overview of Robustness Checks 110

G. SUMMARY 110

V SUBGROUP IMPACTS 113

A. EXPECTED IMPACTS AMONG SUBGROUPS 114

  1. Agency Characteristics 114
  2. Patient Characteristics 116

B. IMPACTS FOR AGENCY SUBGROUPS 118

C. IMPACTS FOR PATIENT SUBGROUPS 126

D. SUMMARY 132

VI COMPARISON BETWEEN DEMONSTRATION YEARS ONE AND TWO 135

A. EXPECTED IMPACTS 135

B. DIFFERENCES AT FOUR AND EIGHT MONTHS FROM ADMISSION 137

C. SUMMARY 146

VII CONCLUSIONS 149

A. KEY FINDINGS 149

B. LIMITATIONS OF THE ANALYSIS 154

CONTENTS *(continued)*

Chapter **Page**

VII *(continued)*

C. POLICY IMPLICATIONS 156

1. Implications of the Reduction in Service Use 156
2. Implications for the Design of a National Program 158

D. SUMMARY 159

REFERENCES 161

APPENDIX A: SUMMARY OF CASE-MIX ADJUSTMENT FOR PAYMENTS DURING THE DEMONSTRATION 163

APPENDIX B: DATA QUALITY 167

APPENDIX C: COEFFICIENT ESTIMATES FROM THE REGRESSION ON TOTAL VISITS PER PATIENT YEAR 179

TABLES

Table		Page
II.1	OUTCOME VARIABLES DESCRIBING MEDICARE SERVICE USE AFTER ADMISSION TO A DEMONSTRATION HOME HEALTH AGENCY	17
II.2	OUTCOME VARIABLES FOR MAIN MULTIVARIATE ANALYSIS, BY SOURCE	18
II.3	ADDITIONAL CONTROL VARIABLES AVAILABLE FROM QUALITY ASSURANCE DATA	23
II.4	PATIENT RECORDS EXCLUDED FROM ANALYSIS SAMPLE, BY TREATMENT STATUS	25
II.5	NUMBER OF PATIENT RECORDS IN ANALYSIS SAMPLE THAT COVER SELECTED TIME PERIODS FROM ADMISSION, BY DEMONSTRATION YEAR AND TREATMENT STATUS	28
II.6	DISTRIBUTION OF ADMISSIONS IN ANALYSIS SAMPLE FOR THE FIRST DEMONSTRATION YEAR, BY TREATMENT STATUS	30
II.7	WEIGHTED MEANS FOR EXPLANATORY VARIABLES, BY TREATMENT STATUS, AND TESTS FOR DIFFERENCES IN TREATMENT AND CONTROL GROUP MEANS	32
III.1	UNITS OF MEASUREMENT FOR KEY OUTCOME VARIABLES IN THE IMPACT ANALYSIS	40
IV.1	CUMULATIVE NUMBER OF VISITS PROVIDED DURING THE YEAR AFTER ADMISSION	70
IV.2	NUMBER OF VISITS PROVIDED DURING THE FIRST 120 DAYS (AT-RISK PERIOD) AFTER ADMISSION, BY SERVICE TYPE	74
IV.3	PROBABILITY OF RECEIVING A GIVEN SERVICE DURING THE FIRST 120 DAYS (AT-RISK PERIOD) AFTER ADMISSION	76

IV.4 NUMBER OF VISITS PROVIDED AFTER THE FIRST 120 DAYS FROM ADMISSION, BY SERVICE TYPE 78

IV.5 PROBABILITY OF RECEIVING A GIVEN SERVICE AFTER THE FIRST 120 DAYS FROM ADMISSION 79

IV.6 TIME FROM ADMISSION UNTIL LAST VISIT (THROUGH ONE YEAR) 82

IV.7 TIMING OF SERVICES FROM ADMISSION 84

IV.8 CHARACTERISTICS OF THE INITIAL PATIENT EPISODE 88

IV.9 IMPACTS OF PROSPECTIVE PAYMENT WITHOUT SAMPLE WEIGHTS 102

IV.10 IMPACTS OF PROSPECTIVE PAYMENT CONTROLLING FOR MORE DETAILED PATIENT-LEVEL CHARACTERISTICS (FROM QUALITY ASSURANCE DATA) 104

IV.11 IMPACTS OF PROSPECTIVE PAYMENT ACCOUNTING FOR OUTLIERS AND CENSORING 106

IV.12 IMPACTS OF PROSPECTIVE PAYMENT BASED ON THE CHANGE FROM PREDEMONSTRATION TO DEMONSTRATION (DIFFERENCE-OF-DIFFERENCE (MODEL) 109

V.1 IMPACT OF PROSPECTIVE PAYMENT ON SERVICE USE, BY WHETHER THE AGENCY HAS A HIGH-USE OR LOW-USE PRACTICE PATTERN 119

V.2 IMPACT OF PROSPECTIVE PAYMENT ON SERVICE USE, BY FOR-PROFIT OR Nonprofit STATUS 120

V.3 IMPACT OF PROSPECTIVE PAYMENT ON SERVICE USE, BY HOSPITAL-BASED OR FREESTANDING 122

V.4 IMPACT OF PROSPECTIVE PAYMENT ON SERVICE USE, BY HOSPITAL-BASED OR FREESTANDING 125

V.5 IMPACT OF PROSPECTIVE PAYMENT ON SERVICE USE, BY UNDER OR OVER (BASE-YEAR) COST LIMITS 127

V.6 IMPACT OF PROSPECTIVE PAYMENT ON SERVICE USE, BY WHETHER PATIENT CAN TAKE ORAL MEDICATIONS INDEPENDENTLY 128

V.7 IMPACT OF PROSPECTIVE PAYMENT ON SERVICE USE, BY WHETHER OTHER CAREGIVERS AVAILABLE 130

V.8 IMPACT OF PROSPECTIVE PAYMENT ON SERVICE USE, BY HIGH OR LOW EXPECTED COSTS 131

FIGURES

Figure **Page**

III.1 AVAILABILITY OF DATA TO COMPARE IMPACTS BETWEEN DEMONSTRATION YEARS 45

IV.1 AVERAGE MONTHLY VISITS DURING THE YEAR AFTER ADMISSION, BY TREATMENT STATUS 67

IV.2 RATE OF SERVICE RECEIPT IN EACH MONTH FROM ADMISSION, BY TREATMENT STATUS 68

IV.3 AVERAGE COMPOSITION OF VISITS IN THE YEAR AFTER ADMISSION, BY TREATMENT STATUS 71

IV.4 LENGTH OF INITIAL PATIENT EPISODE, BY TREATMENT STATUS 92

IV.5 DISTRIBUTION OF SERVICES WITHIN THE INITIAL PATIENT EPISODE, PROSPECTIVELY PAID AGENCIES ONLY 94

IV.6 AVERAGE MONTHLY VISITS AMONG PATIENTS WITH AN ONGOING EPISODE, PROSPECTIVELY PAID AGENCIES ONLY 96

EXECUTIVE SUMMARY

As part of its ongoing effort to study methods of providing more cost-effective care, the Health Care Financing Administration (HCFA) implemented the Per-Episode Home Health Prospective Payment Demonstration. Under the demonstration, participating home health agencies are paid a fixed, lump-sum payment for the first 120 days of each episode of care provided to Medicare beneficiaries and a predetermined rate for each visit thereafter. By allowing agencies to retain most of any surplus payments over cost, this payment method gives agencies an incentive to provide home health care in a cost-efficient manner.

We find that prospective payment *reduced the average number of visits to a patient in the year following admission by 24 percent* compared to their levels under cost-based reimbursement. Reductions in services took place both during and after the 120-day period covered by the episode payment, and they were of a similar proportion for each service type. Prospectively paid agencies achieved these reductions by shortening the overall length of service and by lowering the frequency of visits provided. Reductions occurred among all subgroups of agencies and patients that we investigated, and they were stable between the first and second years of the demonstration. Though further analyses must evaluate the consequences of these declines on patient health and access, non-home health expenditures, and other outcomes before final recommendations can be made, these results strongly suggest that prospective payment is a highly successful method for controlling the costs of Medicare home health care.

BACKGROUND

Ninety-one agencies in five states entered the three-year demonstration at the start of their 1996 fiscal year. Prior to the start of the demonstration, the participating agencies were randomly assigned to either the treatment group or the control group. Agencies assigned to the treatment group have been reimbursed under the demonstration's prospective payment method, while those assigned to the control group continue to be reimbursed under Medicare's method of cost-based reimbursement (the payment method for all home health agencies when the demonstration began).

For each patient admitted to home health care, treatment group agencies receive a lump-sum payment for the first 120 days of home health care, regardless of the number of visits provided or their cost.[1] The amount of the per-episode payment is based on each agency's own costs in the fiscal year immediately preceding its entry into the demonstration, adjusted for changes in its case mix and for inflation. An agency can receive a new per-episode payment for a given Medicare beneficiary only after the 120-day "at-risk" period has ended and a 45-day gap in services has taken place. For each visit after the at-risk period that does not begin a new episode (referred to as the "outlier"

---

[1] Some exceptions apply. For example, durable medical equipment, nonroutine medical supplies, and Part B ambulatory home health services continue to be reimbursed at cost.

period), treatment agencies receive a fixed payment that varies by the type of visit and is again based on their own predemonstration costs.

HCFA reimburses treatment agencies for up to 99 percent of fiscal-year losses up to the (Section 223) cost-per-visit limits that apply to all agencies in the Medicare program. Profits in excess of specified limits must be shared with HCFA.

RESEARCH QUESTIONS AND METHODOLOGY

In this report, we examine the data from the first two years of the demonstration to test various hypotheses about the effects of prospective payment on patients' use of Medicare-covered services. The most important of these hypotheses center on (1) the effect of prospective payment on the number of visits provided in the year after a patient admission, both in total and by type; (2) the effect of prospective payment on the duration of services; (3) whether key impacts of prospective payment differ between selected subgroups of agencies or patients; and (4) whether key impacts changed over the course of the demonstration. In addition, we examined the length, frequency, and other characteristics of the episode of care under a system of prospective payment.

Our main analysis centers on patients admitted to an agency during its first year in the demonstration, because we were able to observe all their subsequent home health use over an extended period of a least 12 months. In a supplementary analysis (based on shorter periods of service use), we then compared key impacts for the first and second demonstration years. The sample for the main analysis included approximately 52,000 home health patients admitted to one of 79 demonstration agencies. (Of the 91 original agencies, 9 were excluded from the analysis because they had incomplete service use data, and 3 left the demonstration near its start.) Medicare claims files provided the data on the outcome variables describing patients' use of services following admission. Data collected at admission for case-mix adjustment and from prior Medicare claims provided measures of preadmission characteristics of patients admitted to agencies participating in the demonstration. Data on agency characteristics were obtained from the agency cost reports and the demonstration implementation contractor.

Ordinary least squares models and logit models were used to estimate program effects, controlling for preexisting differences between treatment and control agencies in patient and agency characteristics. This approach proved crucial to obtaining valid impact estimates, since, despite the randomization of participating agencies, there were several significant differences between treatment and control agencies aside from the method of payment. Observations were weighted so that each agency was represented equally in the analysis. Standard errors of impact estimates were calculated with software designed to account for the effects of sample clustering and weighting, in order to avoid overstating the precision of the estimates. Analyses of the robustness of our regression estimates showed that they were not sensitive to the weighting scheme, statistical methods, or model specifications used. We also found no evidence that our results were affected by incomplete service use data, despite having to drop nine agencies from the analysis for this reason.

FINDINGS

Total Visits Fell an Average of 24 Percent in the Year After Admission

Prospective payment (as defined by the rules of the demonstration) significantly reduced the average number of visits provided to a home health patient in the first 12 months after admission. For patients admitted in year one of the demonstration, control agencies provided an average of about 76 visits, while treatment agencies provided only 58 visits--a decline from prospective payment of about 18 visits, or 24 percent (see Figure 1). More than 90 percent of this decline was due to reductions in visits by skilled nurses and home health aides, a result that is perhaps unsurprising since these services account for the vast majority of visits by most agencies. Skilled nursing visits fell from about 35 to 26, a decline of 26 percent, while home health aide visits fell from about 30 to 23, a decline of 24 percent. Other services also experienced large and significant reductions. Therapy services fell an average of 1.4 visits (19 percent), and medical social worker services fell an average of 0.4 visits (32 percent).

Visits Fell Both During and After the Initial 120-Day Period Covered by the Episode Payment

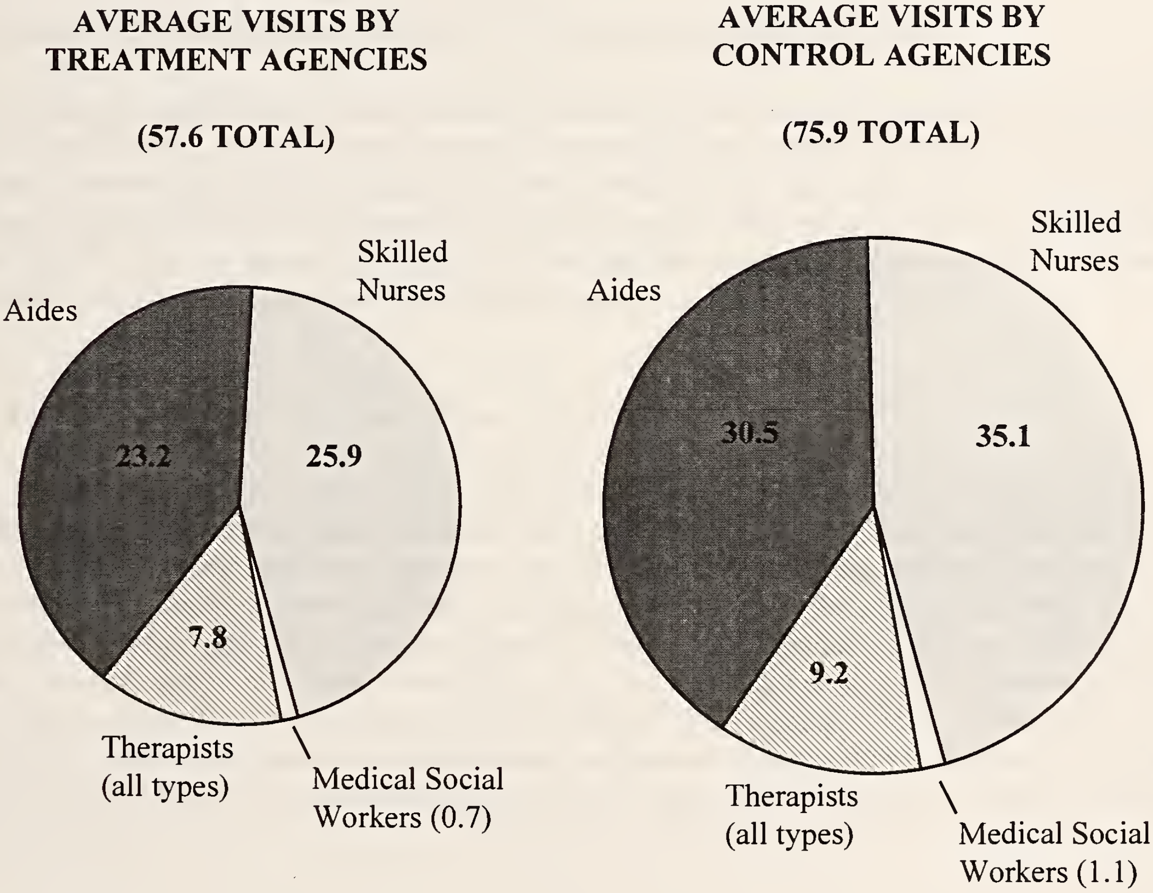
Within the first 120 days after admission (the initial at-risk period), prospective payment reduced the total visits per patient from about 45 to 37, a decline of 17 percent. Skilled nursing services accounted for about half of this decline, falling an average of 4 visits per patient, or 19 percent. Home health aide services fell by close to 3 visits per patient (17 percent), while medical social worker services fell by 0.3 visits (38 percent). The average number of physical, occupational, and speech therapy visits in the at-risk period displayed no significant change from prospective payment. However, the likelihood of receiving services from an occupational therapist fell from about 12 percent to 8 percent. Except for occupational therapy, the likelihood of receiving services of a given type fell either slightly (in the case of skilled nursing) or not at all in the at-risk period.

Following the initial at-risk period, the average number of visits fell by about one-third, from 30 to 20, or about 56 percent of the overall decline in the year after admission. Visits by skilled nurses and home health aides accounted for most of this decline. However, in contrast to the at-risk period, visits by physical therapists and occupational therapists fell significantly, from a combined average of about 2.2 visits to 1.4 visits. Medical social worker visits, which were used rarely after the at-risk period, also fell by a small but significant margin (about 0.1 visits per patient).

Service Lengths Declined Sharply

Earlier discharge was a key reason for all service reductions. Based on episodes observed over a maximum period of 480 days, the average episode length fell from 131 days to 98 days--a drop of more than one month. This decline resulted from broad-based reductions in episode lengths for all types of patients. In total, about three-quarters of the episodes in treatment agencies ended within

FIGURE 1
AVERAGE VISITS PER PATIENT IN THE YEAR AFTER ADMISSION, BY TREATMENT AND CONTROL AGENCIES



SOURCE: Medicare claims data.
NOTE: Treatment group mean is adjusted to account for preexisting differences between treatment and control agencies.

the at-risk period (compared to about two-thirds of those in control agencies), and most ended within 60 days.

Despite the decline in initial episode lengths, treatment agencies did not display a significantly higher rate of second episodes. By 16 months after admission, 10.7 percent of the patients in treatment agencies had experienced multiple episodes of care, compared to 10.2 percent of the patients in control agencies, a statistically insignificant difference.

Frequency of Visits Declined Uniformly Throughout an Episode of Care

Treatment agencies also cut back on the frequency of visits provided, in both the at-risk and the outlier periods. In each month after admission (through at least month 16), patients with an episode of care continuing through the end of the month received an average of one to two visits less from treatment agencies than control agencies. The consistency of this difference through time suggests that agencies did not generally distinguish between the outlier and at-risk periods, but instead pursued initiatives to lower visits at all points after admission.

No Evidence that Agencies "Gamed" the Payment System

We find no evidence to suggest that treatment agencies tried to "game" the payment system in order to obtain additional profit. As noted, the rate of second episodes was fairly small and quite similar for treatment and controls agencies (roughly 10 percent by 16 months post-admission). We also find a very similar rate of episodes with only one or two visits (about four percent), which contradicts the hypothesis that treatment agencies would admit a higher proportion of patients with little or no true care needs. Finally, we find no evidence that treatment agencies increased visits for longer-stay patients in order to obtain additional revenue from visits provided in the outlier period. In fact, the findings suggest that treatment agencies behaved in the outlier period precisely as they did in the at-risk period--by making significant and broad-based reductions in average visits and service length.

Reductions Were Larger Among Agencies with "High-Use" Practice Patterns

Agencies providing a relatively large number of visits per episode prior to the demonstration cut back more visits per patient than relatively "low-use" agencies (about 24 visits per patient in the year after admission for high-use agencies, compared to 12 visits for low-use agencies). These declines were actually quite similar in percentage terms (about 23 percent and 26 percent, respectively), however, because agencies that were relatively high-use prior to demonstration continued to be high-use during the demonstration. The greater reductions by high-use agencies were largely centered in home health aide visits, and they were almost certainly associated with their more limited control of service utilization prior to the demonstration.

Several other agency subgroup differences appear to be associated with practice patterns. For-profit agencies made larger cuts in visits than nonprofits, but the difference was explained entirely by the higher proportion of for-profit agencies with high-use practice patterns prior to the demonstration. Small agencies (those providing fewer than 30,000 visits prior to the demonstration) and freestanding agencies also made greater reductions in visits than their counterparts, but these differences are again at least partly explained by a greater proportion of these agencies having high-use practice patterns prior to the demonstration.

Reductions Were Larger Among Agencies Above the Cost Limits

Agencies above the per-visit cost limits reduced service use about twice as much those below the limits (an average of 31 visits in the year after admission, compared to 15 visits for agencies below the cost limits). This difference can be explained by the very strong incentive for treatment agencies above the limits to generate savings from reduced visits, which is completely in contrast to the incentive faced by control agencies above the limits to lower their per-visit costs through additional visits.

Patient Subgroups Experienced Similar Service Reductions

Patients who had characteristics normally associated with a high level of service use had their average visits reduced by about twice the number of those with relatively low expected service use. However, this difference in impacts (25 visits compared to 13 visits) was only marginally significant, and it was quite similar when measured as a proportion. With respect to other patient subgroups that we evaluated, no significant differences in visit reductions were found, and impacts for each subgroup were generally similar when measured as a proportion.

Impacts Were Stable Between the First and Second Demonstration Years

The available data from the second demonstration year suggest that the impacts of prospective payment remained stable over time. This does not mean, however, that agencies' behavior remained unchanged. To the contrary, both treatment and control agencies made large reductions in visits and service length in the second year (compared to the first), which resulted in no change in the effect of prospective payment between the two years. Even if factors that reduced control agencies' service provision in year two (managed care growth, investigations of fraud and abuse) continued to intensify, they would be unlikely to lead to the dramatic service reductions that treatment agencies achieved.

LIMITATIONS OF THE ANALYSIS

Despite the significance and robustness of our findings, this study does have some important limitations. However, none of these limitations cast any doubt on the validity of the finding that agencies made substantial and broad service reductions in response to prospective payment.

Demonstration Agencies May Have Behaved Differently from Agencies Nationally

A number of possible differences between agencies in and out of the demonstration may lead our findings to overstate or understate the nationwide impacts of prospective payment. Like any study design based on voluntary participation, for example, the agencies in this demonstration may reflect the group most able (or willing) to respond to the incentives of the intervention. If true, our demonstration impacts may overstate the actual reductions in service use that would take place nationally under a similar payment system. In addition, if agencies nationwide had relatively low-use practice patterns compared to demonstration agencies, we would probably overstate the effects of a nationwide program. Of course, if this difference was reversed, our impacts may actually understate these effects.

Despite these concerns, a key factor suggesting that results may generalize well to agencies nationwide is that *all types* of agencies made substantial reductions in their provision of services. For, while agencies of different size, auspice, or practice patterns might be expected to differ both in their capacity and in their desire to respond to demonstration incentives, each agency type made cuts of at least 18 percent in their average visits per patient in the year after admission.

A Demonstration Payment System Would Likely Differ from a National Payment System

A second limitation, related to the issue of generalizability, is that the national program of prospective payment (begun in October 2000) differs from the one implemented for the demonstration. Under the national program, for example, agencies are not protected from incurring financial losses and do not have their payment based on their own prior cost per episode. These and other differences may lead at least some agencies to act quite differently under the national payment system than they have under the demonstration. (For example, when faced with the potential for real losses and a payment not based on their own historical costs, some agencies may need to reduce visits even further under the national system to remain solvent.) Whether, and how, agency behavior would differ under the national payment system is uncertain. However, there is no reason to believe that the large-scale reductions in service use observed under the demonstration would not exist under the national system of per-episode prospective payment as well.

Findings Are Limited to the Effects on Service Use

This report does not provide information on the consequences of reducing services. In a series of additional reports, we address this issue by examining the effects of the demonstration on the quality of care and patient outcomes, patient selection and retention, and the utilization of other Medicare and non-Medicare reimbursed services. In addition, we examine the impacts of the demonstration on agencies' costs per visit and costs per episode, as well as their profits and losses. A final report will combine the findings across outcome measures.

Findings Are Based on Only Two Years of Home Health Admissions

Because this report examined the effects of prospective payment over just two years of the (three-year) demonstration, our findings may not accurately reflect the long-term impacts of prospective payment. Within just the first year of the demonstration, however, nearly all participating agencies in the treatment group made dramatic service use cuts that may be directly attributed to the new system of payment. A key hypothesis that the demonstration was designed to investigate--whether prospective payment could curtail the tremendous growth in home health services--was thus strongly affirmed, despite the limited period of investigation.

POLICY IMPLICATIONS

The findings from this report provide key information on the effects of prospective payment at a time when HCFA, under mandate from the 1997 Balanced Budget Act, must implement a nationwide prospective payment system. Two basic results from this study are perhaps most relevant to HCFA's efforts. First, prospective payment generates sizable reductions in service use, which may in turn lead to substantial savings to Medicare. Second, the specific version of prospective payment evaluated for the demonstration achieved precisely the desired effect, suggesting that it provides a sound starting point for implementing a national payment system.

Prospective Payment Offers a Substantial Opportunity for Savings

The large service reductions by treatment agencies suggest that a nationwide prospective payment system may lead to substantial savings relative to a system of cost-based reimbursement. However, such savings may be somewhat smaller proportionally than the overall reduction in visits (roughly 24 percent), because agencies' cost per visit may increase under prospective payment (Cheh and Trenholm 1998). Nonetheless, even if savings were only a fraction of this size, HCFA may have the opportunity to reduce spending on home health care while still allowing agencies to prosper.

While similar cost savings might be achieved through direct regulation of the home health industry, prospective payment is likely to be a better approach. Consider, for example, the differences between a system of prospective payment and the recently enacted interim payment system (IPS).[2] Although the IPS has led to sharp reductions in Medicare home health expenditures, these reductions have been achieved through caps imposed on per-beneficiary spending and per-visit costs. Effectively, these caps limit not only the level of care provision but also the mode of savings, since some approaches to cost-cutting may not benefit the agency. (For example, an agency would not want to generate savings through approaches that might drive its per-visit costs above the cap.) In contrast, a per-episode prospective payment system offers great flexibility, both in the way

---

[2] The 1997 Balanced Budget Act mandated the IPS until the date that HCFA has implemented prospective payment. Under IPS, agencies received the lowest of (1) their actual allowable costs; (2) their aggregate cost per visit, lowered to 105 percent of the established Section 223 Limits for the region; and (3) a new aggregate per-beneficiary limit.

agencies serve their patients and how they generate savings. Moreover, because agencies` profits rise as their costs decline, prospective payment creates the incentive to generate as much savings as possible (short of degrading care quality). Prospective payment thus provides an opportunity for savings that is similar to what might be imposed under a more regulated system of payment, while allowing agencies more flexibility to provide care that is tailored to the needs of individual patients.

Demonstration Offers a Sound Starting Point for Design of a National Payment System

Although a voluntary demonstration cannot address all the issues that must be considered in the design of a national payment system, the basic demonstration design--a 120-day at-risk period followed by prospective per-visit payments--appears to be a promising approach. Indeed, based on the complete set of study findings to date, the demonstration design may be capable of generating large savings relative to a system of cost-based reimbursement while maintaining both access to care and care quality. The demonstration design thus provides a sound starting point for implementing a national payment system, which can then be modified as necessary to address concerns that were beyond the scope of the demonstration, including service availability, regional variation in service use and costs, and long-term efficacy and operability.

I. THE PER-EPISODE HOME HEALTH DEMONSTRATION AND EVALUATION

The Per-Episode Home Health Prospective Payment Demonstration of the Health Care Financing Administration (HCFA) tests the extent to which prospective payment for Medicare home health services increases efficiency in service provision. Per-episode payment encourages efficiency by giving agencies the financial incentive to reduce their costs. Specifically, under the demonstration payment system, any savings generated from lower costs per episode of patient care might result in agency profit.[1] These incentives differ greatly from those under a system of cost-based reimbursement, which does not provide a reward for cost containment.

In this report, we present findings from an analysis of demonstration impacts on patients' use of Medicare home health services during agencies' first two years of participation in the demonstration.[2] These impacts are critical because a reduction in home health use is the key mechanism through which agencies may lower costs per patient episode. In addition, any reductions in home health use may have important consequences for the quality of care received and the use of other Medicare-covered services.

The remainder of this chapter provides an overview of the Medicare home health benefit and the Medicare-certified home health industry and describes the Per-Episode Home Health Prospective

---

[1] Strictly speaking, only for-profit agencies earn profits; nonprofit agencies generate surpluses that cannot be distributed to "owners" of the organization. A for-profit firm can distribute profits (or surpluses) as it wishes, including bonuses to employees, distributions to shareholders, and so forth. However, for brevity, we use the term "profits" in this report to refer to surpluses generated by nonprofit agencies, as well as profits earned by for-profit agencies.

[2] Originally, this report was to focus on use during the entire three-year period of the demonstration. However, Congress mandated that by October 2000, the Medicare program will use a prospective payment system when paying for home health services. In turn, HCFA requested that this report be submitted before the demonstration has been completed, so the findings can be considered during the design of the new payment system.

Payment Demonstration. A reader who is familiar with Medicare home health may wish to proceed directly to Section B, which describes the demonstration procedures.

A. THE MEDICARE HOME HEALTH BENEFIT

Congress established the Medicare home health care benefit in 1965, when the original Medicare program was created. Home health benefits were included to offer beneficiaries with acute conditions a less intensive and less expensive alternative to inpatient hospital care. At different times since the inception of the Medicare program, the home health benefit has been modified, partly to increase access to care. The current benefit covers home health services mostly under Part A; neither a deductible nor coinsurance applies.[3] To be eligible for home health benefits, the beneficiary must (1) have Medicare coverage; (2) be homebound; (3) be under the care of a physician; and (4) need skilled nursing, physical therapy, or speech therapy on a part-time or intermittent basis.[4]

Coverage under the home health benefit broadened considerably after the settlement of a lawsuit against HCFA in 1989, which, in turn, contributed to dramatic growth in Medicare home health

---

[3] In some rare instances, such as when an individual does not qualify as homebound or does not have Part A coverage, home health care may be covered under Medicare Part B. In cases of ambulatory Part B home health, deductibles and copayments do apply. In addition, under the Balanced Budget Act of 1997, the long-term use of home health services was transferred to Part B. In these cases, there are no deductibles or copayments. Our analysis period predates this change, however.

[4] Skilled nursing services are covered as long as (1) a physician has ordered them, (2) they are required on a part-time or intermittent basis, (3) they require the skills of a registered nurse (or of a licensed practical nurse or licensed vocational nurse under a registered nurse's supervision), and (4) they are reasonable and necessary to treat an illness or injury. Physical therapy and speech therapy are covered if a physician's assessment recommends these services. Beneficiaries who need only occupational therapy are entitled to benefits only if they have established a prior need for skilled nursing care, speech therapy, or physical therapy in the current or previous certification period (see Teplitsky and Janson 1985-1992, p. VII.23, Section 204.4).

expenditures after that time. Medicare payments for home health care rose from $2.4 billion in 1989 to $16.8 billion in 1997, more than tripling the share of total Medicare spending on home health (Health Care Financing Review 1998). Nearly all this growth has been due to an increase in the provision of services, which has coincided with a dramatic expansion in industry size. Between 1989 and 1997, the number of home health users per 1,000 beneficiaries rose from 51 to 109, and the average number of visits per user rose from 27 to 73 (Health Care Financing Review 1998; U.S. General Accounting Office 2000).[5] Coinciding with this growth, the number of Medicare-certified home health agencies rose from 5,700 to more than 10,000 (U.S. General Accounting Office 1998). Service use patterns displayed striking differences across regions throughout this period. In 1997, for example, the average number of visits per user ranged from 37 visits in Maryland to 161 visits in Louisiana (U.S. General Accounting Office 2000).

The dramatic growth in Medicare home health expenditures, combined with striking regional variation in service use and recent investigations of industry fraud and abuse, prompted Congress to legislate changes to the Medicare home health benefit as part of the Balanced Budget Act of 1997. The Act mandated the implementation of a per-episode prospective payment system for Medicare home health by 1999. This was later changed to October 1, 2000 under the Omnibus Consolidated Emergency Supplemental Appropriations Act of 1999. Other changes to the home health benefit under the Balanced Budget Act of 1997 include:

---

[5] These figures reflect only Medicare beneficiaries in fee-for-service.

- Lowering the per-visit cost limit (known as the Section 223 limits) from 112 percent of the mean to 105 percent of the median cost for freestanding agencies in the region. The Section 223 limits were frozen for reporting periods that began between July 1, 1994, and June 30, 1996.

- Defining the maximum payment for an agency by using a new algorithm that is based on annual per-beneficiary costs or per-visit costs in a base year, whichever is lower. This algorithm is commonly referred to as the Interim Payment System (IPS), because it is intended to last only until the mandated prospective payment system takes effect.

- Eliminating coverage for blood drawing when it is the only home health service required

- Redefining "part-time" and "intermittent" care

- Redefining service location on the basis of the location of the patient rather than of the agency

- Requiring additional billing information (an identifier for admitting physician and visit length)

The changes in the per-visit cost limit and the algorithm for payment (the so-called IPS system) have no effect on demonstration agencies until after their participation in the demonstration has ended. However, demonstration agencies are affected by the other changes.

B. THE PER-EPISODE DEMONSTRATION

Because a cost-reimbursed payment system does not provide any mechanism that enables home health agencies to earn profits by cutting costs, it offers no incentive for the agencies to provide services efficiently. Moreover, by reimbursing costs up to allowable limits, the system effectively subsidizes inefficient providers. The intent of per-episode prospective payment is to increase efficiency, by using the opportunity to generate profits as the primary incentive.

Ninety-one Medicare-certified home health agencies in five states--California, Florida, Illinois, Massachusetts, and Texas--enrolled in the three-year per-episode demonstration.[6] Of these, 47 were randomly assigned to the treatment group and received per-episode payment. The remaining 44 were assigned to the control group and continued to operate under cost reimbursement. One control agency then switched into the treatment group near the start of the demonstration, leading to a revised total of 48 treatment agencies and 43 control agencies.[7] Each agency entered the demonstration and began implementing prospective payment at the start of its fiscal year. The first entrants to the demonstration began implementing prospective payment in June 1995; the latest entrants began in January 1996. Demonstration operations were planned to continue through December 1998. (The demonstration was extended for treatment group agencies until prospective payment Medicare home health care is implemented nationally.)

Mathematica Policy Research, Inc. is the contractor responsible for the process and impacts evaluation of the demonstration. Several other organizations are also participating in the demonstration. Abt Associates, Inc. is the implementation contractor responsible for recruiting demonstration agencies, monitoring the status of demonstration operations, and calculating certain statistics required to determine agency payments. Palmetto Government Benefits Administrator (PGBA) is the fiscal intermediary (FI) responsible for review of claims and agency payment for both treatment and control agencies. The Center for Health Policy Research (CHPR) at the University

---

[6] Reflecting the United States more generally, considerable variation existed in the use of Medicare home health across the five demonstration states. In 1997, the average visits provided per beneficiary were as follows: California, 47; Illinois, 47; Florida, 70; Massachusetts, 89; and Texas, 134 (U.S. General Accounting Office 1998). In a future report, we will compare the demonstration agencies with those nationwide.

[7] The switch was made at the request of an agency that had established a network with two other agencies assigned to the treatment group. The three agencies plan to merge fully at the end of the demonstration.

of Colorado is responsible for designing and implementing a quality assurance (QA) system for the demonstration agencies.

1. Demonstration Payment and Incentives

HCFA developed the Home Health Prospective Payment Demonstration to assess whether the profit motive can increase efficiency in providing Medicare home health care and thereby reduce public expenditures, without sacrificing access to care or the quality of care. Phase I of the demonstration, which tested per-visit prospective rate setting, gave agencies an opportunity to generate profits (and avoid losses) by reducing per-visit costs.[8] The evaluation of Phase I found that some agencies earned profits but did little to contain overall costs, because they had no incentive to limit the services they provided. The current phase of the demonstration, Phase II, tests a per-episode prospective payment system. This system encourages agencies to reduce their per-visit costs, and it provides the additional incentive to reduce visits per patient episode. Thus, the opportunity for cost savings is much greater under the Phase II payment system than under the Phase I system.

a. Payment

Agencies selected for the treatment group receive a lump-sum payment for the first 120 days of home health care, regardless of the number or cost of visits provided.[9] The agencies are thus "at

---

[8] Although the per-visit demonstration was implemented in the same five states, most of the agencies participating in the per-episode demonstration did not participate in the per-visit demonstration. (Only agencies in the per-visit control group were eligible to do so. For details on the results of the per-visit demonstration, see Brown et al. [1995].)

[9] Durable medical equipment, nonroutine medical supplies, and Part B ambulatory home health services continue to be reimbursed at cost throughout the demonstration. In addition, if an agency did not provide one or more of the six Medicare services during the base year but begins to do so

(continued...)

risk" for the costs of care incurred during this period. Only after the 120-day at-risk period and after a 45-day gap in services have elapsed can an agency receive a new per-episode payment for a given Medicare beneficiary. For each visit beyond 120 days that does not begin a new episode (referred to as "the outlier period"), treatment agencies receive a fixed payment rate that varies by the type of visit. In the demonstration, a treatment agency is also paid on a per-visit basis for visits made to patients admitted before the agency began demonstration operations ("phase-in" visits), and to patients admitted within 120 days of the end of demonstration operations in that agency ("phase-out" visits).

Agencies selected for the control group receive payments based on the cost-based system in place at the start of the demonstration. Specifically, the control agencies are reimbursed for their actual per-visit costs, up to 112 percent of the mean cost incurred by all agencies (for the agency's mix of visits) in the same geographic area.

b. Prospective Rate Setting

Prospective (per-episode) rates for the at-risk period are based on a treatment agency's costs and episode profile in the fiscal year preceding its entry into the demonstration (the base year), adjusted for inflation and changes in case mix in each evaluation year.[10] The episode profile is the average number of visits provided by the agency during an episode, calculated for each of the six types of visits covered by Medicare. Payments for outlier, phase-in, and phase-out visits are also based on

---

[9] (...continued)
during the demonstration, then those visits are also reimbursed at cost during the demonstration, as are the costs of care for which Medicare is a secondary payer.

[10] For more information on payment rates under the demonstration, see Phillips et al. (1995).

the agency's base-year per-visit costs (adjusted for inflation).[11] HCFA's market basket is used to adjust both the per-visit and per-episode rates for inflation. Both the per-episode and the per-visit payment rates were subject to HCFA's statutory home health cost limits.

The case-mix adjuster classifies each patient into 1 of 18 groups on the basis of 12 variables that describe the patient's characteristics and hospital stay history. From this information, an aggregate case-mix index is created for each agency. At the end of each year of the demonstration, an agency's case-mix index for that year is compared with its case-mix index for the base quarter (the last quarter of the base year). If the agency's case mix differs from its base-quarter case mix, then its aggregate payment is retrospectively adjusted (see Appendix A for additional details).

c. Profit Sharing and Loss Sharing

To counteract the incentive to dramatically reduce services at the expense of quality, as well as to prevent agencies from realizing windfall profits at public expense, HCFA shares in any profits above selected levels. If the total of a treatment agency's per-episode and per-visit prospective payments is greater than the costs for services covered by these payments, then any resulting profit greater than five percent of total allowable costs for these services is subject to profit sharing. HCFA's share is 25 percent for profits between 5 and 15 percent of total allowable costs; the share increases for profit rates greater than 15 percent by an amount that varies with the demonstration year.

HCFA has provided a loss-sharing arrangement as a means of encouraging agencies to participate in the demonstration. It reimburses treatment agencies for 99 percent of losses occurring

---

[11] Because complete data for episode profiles and settled cost reports are not available for a given year until some months after that year has ended, the initial lump-sum and per-visit rates used in the demonstration were preliminary and were revised after final base-year data became available.

during the first demonstration year, and for 98 and 97 percent of losses during the second and third demonstration years, respectively, as long as total payments are within the demonstration cost limits.

2. Other Demonstration Procedures

Three other important demonstration procedures involve the process of medical review, the methods for billing, and the requirements for QA. Treatment agencies are strongly affected by all these demonstration procedures; control agencies are principally affected by the QA requirements. In this section, we briefly discuss each procedure.

a. Medical Review

For agencies in the treatment group, PGBA conducts a limited medical review (known as an "abbreviated" medical review) of the care delivered during the at-risk period of the patient episode. This review, which is limited to the admission bill, seeks to determine whether the patient met the coverage criteria for home health care, and whether the agency provided at least one visit that met these criteria. As a condition of payment, PGBA requires the agency to submit either HCFA 485 and 486 forms (which contain information on the patient's health and eligibility status, as well as on the home health plan of treatment) or clinical notes for admissions that coincide with an episode eligible for prospective payment. The medical review is based on these materials. All services paid for under per-visit rate setting are subject to the usual medical review, under which a sample of claims is reviewed to ensure that each visit is medically reasonable and necessary.

Although all episodes initially were subject to abbreviated medical review, Medicare reduced the proportion to 25 percent in May 1996. It made this change because the claims reviews took

longer than expected, and because 100 percent medical review was considered unrealistic under a national program.

Medical review for control agencies continues under the current (nondemonstration) regulations. For some control agencies, however, PGBA's review procedures may differ from those of their predemonstration FIs, requiring them to adopt some minor procedural changes.

b. Billing

Treatment agencies must submit an admission bill to PGBA in order to initiate an episode of care. If a treatment admission claim is accepted (subsequent to abbreviated medical review), then the per-episode payment is provided as a lump sum.[12] Treatment agencies also must submit interim bills for services provided during the remainder of the at-risk period, although they receive payments only for supplies.[13] For services provided near the start of the outlier period, agencies must "split" its standard bill into two bills--one for services provided up to the end of the at-risk period, and one for services provided from the start of the outlier period.

When a patient is discharged during either the at-risk period or the outlier period, agencies must submit a discharge bill to terminate the episode. PGBA will not initiate a new episode for a given patient unless a previous episode has been terminated. In addition, before initiating a new episode, PGBA determines that the 120-day risk period and a 45-day gap have elapsed.

---

[12] If the admission claim is denied, interim claims for that episode are suspended for 65 days to await appeal. If an appeal is filed, interim claims are suspended until a decision on the appeal for the admission claim is made. If an admission claim is denied but an appeal is not filed within 65 days, or if the denial of the admission claim is upheld on appeal, then suspended interim claims are released for possible payment under the agency's per-visit rates.

[13] Agencies also must submit interim bills to receive payment during outlier periods, and to correctly calculate costs for profit or loss sharing with HCFA.

Periodic interim payments (PIPs), a means of "smoothing" cash flow for home health agencies, were originally discontinued for treatment agencies. However, a similar periodic payment system (termed "biweekly interim payments") was later introduced by PGBA to address cash flow problems in some agencies.

Control agencies continue to submit bills under cost-reimbursement and continue to be eligible for PIPs. PGBA bases the amount of the PIP on the agency's average cost for each type of visit. For some agencies, this method differed from the practices of their predemonstration FIs, which based the PIP on overall agency average cost per visit. The new method may have led to minor differences in the PIPs some control agencies received before the demonstration compared with those received after they entered the demonstration.

c. Quality Assurance

All agencies participating in the demonstration (whether in the treatment or control group) are required to collect and submit patient-specific information to CHPR, the demonstration QA contractor. The QA procedures follow a continuous quality improvement approach. Visiting staff from demonstration agencies are required to collect information (primarily on functional status and medical condition) at admission and at discharge, or 120 days after admission, whichever comes first. Similar information is also collected before admission to an inpatient facility for stays of 48 hours or longer, and on a patient's return to home health care after such an inpatient stay. CHPR uses this information to develop profiles describing patient outcomes at each agency. These profiles are given to the demonstration agencies to help them improve the quality of care they provide.

C. GUIDE TO THE REST OF THIS REPORT

The next chapter of this report describes the data sources, samples, and variables used to examine home health service use under the demonstration. Chapter III describes the statistical methods used to measure demonstration impacts. In Chapter IV, we present our hypotheses and findings from an analysis of demonstration impacts on service use during the first year after patient admission. We also provide descriptive data on service use during the at-risk period and (one-year) outlier period of the patient episode. Chapter V presents demonstration impacts for specific agency and patient subgroups, and Chapter VI reports on any changes in impacts between the first and second demonstration years. In Chapter VII, we summarize the key findings and present our conclusions.

II. DATA

Our study of home health use under the demonstration relies on a main data file containing 119,358 records for episodes of care beginning in the 91 demonstration home health agencies between their demonstration start and August 31, 1997. The file is primarily built from Medicare claims data extracted from the standard analytic files in May 1998. Allowing a four-month window for the standard analytic file to include all paid claims, our file should contain complete service use data through December 31, 1997.[1]

Because agencies entered the demonstration at the start of their 1996 fiscal years, the period covered by the file varies by agency. A majority of the agencies--57--entered the demonstration on December 31, 1995, enabling us to identify 20 months of demonstration admissions for them. The remaining 34 agencies enrolled between May 30, 1995, and November 31, 1995, for identification of 21 to 27 months of demonstration admissions.

Each record in the main file includes the following information: (1) the patient's use of Medicare-covered services from the time of home health admission (the episode start) through a maximum of 480 days[2]; (2) a number of patient characteristics at admission; (3) the patient's Medicare service use from the time preceding admission; (4) the predemonstration characteristics

---

[1] Although the vast majority of claims should be in the standard analytic file within the four months, a small number of demonstration agencies experienced significant delays in billing that may have caused the file to undercount the number of visits provided. In general, these agencies were dropped from our analysis. See Appendix B, Data Quality, for further information on this topic.

[2] Because our service use data are complete only through December 31, 1997, any episodes starting after September 7, 1996, provide complete information on fewer than 480 days of use. The most severely truncated records are for admissions occurring on our cutoff date, August 31, 1997, which include home health use during the first 122 days from admission.

of the home health agency providing care; and (5) the characteristics of the area in which the agency is located.

From this main file, we created an analysis sample of 86,448 patient records in 79 agencies that were appropriate for evaluating demonstration impacts. We had two reasons for excluding records from the analysis sample. First, we excluded all the records for 12 demonstration agencies that had left the demonstration or had indicated delays in billing that could have led to large undercounts of service use. Second, we excluded a number of individual records for several reasons associated with missing or invalid data.

In the next two sections of this chapter, we provide further details on the construction of the main file and the methodology used to arrive at the analysis sample. In the last section, we summarize the control variables that will be used in the impact analysis to account for possible preexisting differences between the treatment and control groups. (For a detailed discussion of the statistical methods used to conduct our analysis, see Chapter III.)

A. CONSTRUCTING THE MAIN DATA FILE

Each patient record included a substantial number of variables from Medicare claims data and a limited number of variables from other sources. Medicare claims data gave us the information to construct measures of home health utilization. We also used these data to construct several control variables measuring preadmission use of Medicare services. The Enrollment Database (EDB) file of the Health Care Financing Administration (HCFA) provided basic demographic information on patients, including their race/ethnicity, gender, and age. Demonstration data collected for case-mix adjustment and for evaluation of quality of care provided more detailed information on patient health status at admission. In addition, case-mix and claims data pertaining to the predemonstration period were used to construct a measure of the "practice patterns" of individual agencies. Agency cost

reports for the base year and county-level data from the Area Resource File (ARF) provided the remaining data required to construct agency-level measures.

1. Identifying Individual Patient Episodes (Records)

For patients in the treatment and control agencies, we used data from UB-92 bill record files obtained from the demonstration fiscal intermediary, Palmetto Government Benefits Administrator (PGBA), to identify home health episodes as defined by demonstration rules.[3] We scanned the UB-92 files, beginning the file search from the date of each agency's enrollment in the demonstration, to identify the first admission for an individual and the complete set of that person's subsequent bill records. The first demonstration admission determined the episode start date from which we constructed the first record for the patient. To determine the end of the initial episode and the start of any subsequent episodes, we tracked bill records until we observed a 45-day gap in care that began after the end of the at-risk period (that is, after the first 120 days from admission). This procedure was followed regardless of whether the agency discharged and readmitted a patient within the first 165 (120 + 45) days after the initial admission.

If we observed a readmission for a patient after 165 days, and a 45-day gap in care had taken place, we created a second record for the patient corresponding to this second demonstration episode. We then repeated the process until we had constructed, for each patient, a series of records for all the episodes beginning in the agency between its demonstration start and August 31, 1997.

---

[3] We used the UB-92 bill records rather than HCFA claims data principally because they contain data on patient characteristics at admission. These data were then used to construct several key control variables.

2. Creating Outcome Variables

To construct the outcome variables, we used claims data from the standard analytic file created by HCFA.[4] After matching the episodes identified from UB-92 bill record data to the standard analytic files, we extracted all claims for a given patient indicating that the admitting agency had provided home health care during the 16 months (480 days) after the episode start. Both Part A and Part B home health data were extracted, but only Part B claims subject to a per-episode payment were included in the outcome measures.[5]

The outcome variables that we constructed from this data, listed in Table II.1, fall into three broad categories: (1) measures of postadmission service use, overall and by type; (2) measures of the timing and duration of these services; and (3) measures of the characteristics of the patient episode. As noted in the table, we measure these variables in several different ways. (The section on Units of Analysis in Chapter III provides additional details.)

3. Creating Control Variables

As discussed in Chapter III, the impact analysis relies on regression models that control for possible preexisting differences between treatment and control agencies. The control variables used in the main regression models, shown in Table II.2, fall into one of two general categories: (1) patient characteristics, including basic demographics at admission and Medicare service use during

---

[4] We chose to use HCFA's standard analytic file because it includes adjustments for any voided or amended bills. Although we could have used the UB-92 data and mimicked HCFA's adjustment algorithms, we could not have been sure that we would replicate all the adjustments perfectly.

[5] Under demonstration procedures, the only Part B claims that fall under a per-episode payment are those for patients with no Part A coverage who received home health care that would have otherwise been covered by Part A.

TABLE II.1

OUTCOME VARIABLES DESCRIBING MEDICARE SERVICE USE AFTER ADMISSION TO A DEMONSTRATION HOME HEALTH AGENCY

Service Use[a]

Total Number of Visits
Probability of a Visit (Any Type)
Number of Skilled Nursing Visits
Number of Home Health Aide Visits
Number of Physical Therapy Visits
Number of Occupational Therapy Visits
Number of Speech Therapy Visits
Number of Medical Social Worker Visits
Probability of a Visit (Each Type)

Timing and Duration of Services

Days Until Last Visit[b]
Probability of Last Visit, by Selected Periods After Admission[b]
Number of Visits in the First Month, by Selected Types
Probability of a Medical Social Worker Visit in the First Month

Characteristics of the Patient Episode[c]

Visits During the At-Risk Period, by Type
Visits During the Outlier Period, by Type
Length of the Episode (Days)
Probability of a Second Episode
Probability of Only One or Two Visits

---

[a] Measured from admission through one year, as well as the following three periods within the year: (1) months 1 though 4 (the at-risk period), (2) months 5 through 8, and (3) months 9 through 12. Total visits and the likelihood of any visit also are measured monthly.

[b] Measured over a maximum of 365 days. Patients remaining in a care episode on day 365 are assigned a censored value of 366 days.

[c] Measured over a maximum of 480 days. Patients continuing an episode past day 480 are assigned a censored value of 481 days.

TABLE II.2

CONTROL VARIABLES FOR MAIN MULTIVARIATE ANALYSIS, BY SOURCE

Patient Level		Agency Level		Area Level
Patient Characteristics at Admission[a]	Medicare Service Use Prior to Admission[b]	Base-Quarter Patient Service Use[c]	Agency Characteristics[d]	Area Characteristics[e]
Age	Length of preadmission inpatient stay (if any)	Agency practice-pattern index	Chain member	Physicians per 10,000 (1994)
Gender	Whether in skilled nursing facility within 14 days before admission		Hospital based	Nursing home beds per 100 elderly residents (1991)
Race	Number of home health visits in six months prior to admission		Proprietary	Hospital occupancy rate (1993)
Original reason for entitlement	Total Part A Medicare reimbursement in six months prior to admission		Small agency (<30,000 visits in base year)	
Whether has cancer			State	
Whether has diabetes			Rural	
Whether has decubiti				
Whether needs complex wound care				
Whether has limitations in activities of daily living				
Whether admitted to home health from hospital				
Whether has Medicaid				
Date of admission				

[a] From Medicare Enrollment Database; UB-92 remarks.

[b] From Medicare standard analytic files.

[c] From Abt base-quarter case-mix file.

[d] From base-year cost reports and Abt enrollment file.

[e] From Area Resource File.

the six months preceding admission; and (2) characteristics of the admitting agency and area in which the agency resides.

a. Patient Characteristics

Variables on patient characteristics control for possible case-mix differences between treatment and control agencies. We expect that, compared with their counterparts, individuals who are more severely ill, have diagnoses requiring greater amounts of care, and have greater limitations in Activities of Daily Living (ADLs) will require more Medicare home health services.

We obtained data on patient characteristics as they were at the start of the home health episode from three sources: (1) the UB-92 Medicare billing form, which includes data collected for the demonstration's case-mix adjuster; (2) Medicare enrollment databases; and (3) Medicare standard analytic files. In the remarks field of the admission bill, all agencies were required to submit information on patient characteristics necessary to create the 18-category Home Health Utilization Group case-mix adjuster. These characteristics include measures of impairment in ADLs and whether the patient has certain medical conditions (cancer, diabetes, decubiti) and care needs (complex wound care). Medicare enrollment files provided us with basic patient demographic information, including the patient's age (at the start of home health episode), gender, race/ethnicity, and disability status (from the original reason for Medicare qualification).

From the standard analytic files, we constructed measures of Medicare service use for the period preceding the home health admission to capture the patients' severity of illness. These measures included indicators of recent acute illness (whether admitted from the hospital, length of previous hospital stay) and longer-term home health use (covering the six months preceding admission). We

assigned the sample mean in the case of patients with incomplete data for that period, and created a set of three indicator variables to identify the reason for any imputation.[6]

b. Agency and Area Baseline Characteristics

We constructed a number of agency-level control variables because different types of agencies might have varying abilities and incentives to reduce visits in response to the demonstration. In addition, certain types of agencies might serve a mix of patients requiring more (or less) care than the patients of other agencies. For example, proprietary and nonprofit agencies might have different preexisting practice patterns with respect to the number of visits rendered per episode, and, relative to freestanding agencies, hospital-based agencies might serve patients who are somewhat more-acutely ill.

Most of the agency-level variables are based on data from agencies' predemonstration year ("base-year") Medicare cost reports and from the demonstration implementation contractor, Abt Associates (Abt). The cost reports supplied data on the agencies' state, for-profit status, and affiliation; they also enabled us to measure the agency size, by using the total number of visits provided during the base year.[7] Data collected by Abt before the demonstration start provided information on whether an agency was a member of a chain or was located in a rural area (as defined by the U.S. Bureau of the Census).

---

[6] The indicator variables identified (1) patients in a health maintenance organization (HMO) during the six-month preadmission period; (2) Medicare as a secondary payer during that period; and (3) patients younger than age 65.5, who generally were not enrolled in Medicare for a full six months. For this last group, we assigned the mean values for those aged 65.5 to 66 years, rather than the mean for the entire sample.

[7] We defined small agencies as those providing fewer than 30,000 visits during the base year (approximately 25 percent of the sample), and large agencies as providing 30,000 visits or more.

The agency practice-pattern variable is an index of the average number of visits per episode provided by an agency during the first 120 days of base-quarter episodes relative to the average number provided by other demonstration agencies. The index accounts for differences across agencies in the average number of visits of each type and in the characteristics of their patients.[8] A value greater than one indicates that, controlling for differences in case mix, an agency provided more visits during the 120-day period than did other demonstration agencies during the quarter preceding the demonstration.

Selected control variables were used to define the following four pairs of agency-level subgroups: (1) an agency's practice pattern (low use or high use), (2) its proprietary status (for profit or nonprofit), (3) its size (small or large), and (4) its auspice (freestanding or hospital-based). In addition, we used data from the cost reports submitted for the base year to define a fifth agency-level subgroup--whether an agency was above or below its base-year cost limits.

We also controlled for differences in area-level characteristics that might influence the care an agency renders. The area-level control variables that we constructed included the number of nursing

---

[8] The specific construction is as follows. Let subscript i refer to the service type and subscript j refer to case-mix cell. Using the case-mix adjuster (developed by Abt), we classify an agency's patients into 1 of 18 case-mix cells (j=1,..,18). Within each case-mix cell, we multiply the average visits of each type for a given agency (n_{ij}) by its national cost limit in the base year (w_i), and then sum across the visit types. This sum essentially reflects a weighted count of the average visits for an agency within a case-mix cell. We then use these weighted counts to construct the ratio of the agency's average number of visits received by patients in the case-mix cell $\Sigma w_i n_{ij}$ to the average among all agencies for this cell ($\Sigma w_i N_{ij}$). Finally, for each agency, we arrive at the practice-pattern index by summing across the 18 case-mix ratios, weighting each ratio by the agency's proportion of episodes in the case-mix cell during the base quarter (p_j). Thus, for each agency, the index practice pattern is given by:

$$\sum_j p_j \left(\frac{\sum_i w_i n_{ij}}{\sum_i w_i N_{ij}} \right).$$

home beds, the number of physicians, and the hospital occupancy rate. All these variables were constructed at the county level, using data from the Area Resource File (ARF).

4. Additional Control Variables

In a set of limited analyses designed to investigate the robustness of our results, we used an additional set of control variables that contain more detailed case-mix information on patients. These variables, shown in Table II.3, were constructed from the demonstration quality assurance (QA) data, collected by the Center for Health Policy Research at the University of Colorado. We used these variables in only a limited capacity because they were unavailable for more than 40 percent of the patients in our sample.[9]

Selected QA variables were used to define three pairs of patient-level subgroups: (1) patients who can or cannot take oral medications independently, (2) patients who do or do not have people available to provide assistance, and (3) patients whose expected costs are relatively high or relatively low. The last subgroup was constructed using a number of QA variables identifying the agency's planned treatments. These variables are (1) administration of insulin, (2) administration of other intramuscular or subcutaneous injections, (3) management and evaluation of patient care plan, (4) dysphagia treatment, (5) partial/complete bedbath, and (6) personal care.[10] The remaining variables used to construct this subgroup were obtained from the remarks field of the UB-92 as well as the QA

---

[9] QA data for much of the sample were missing for two reasons. First, agencies did not collect these data until May 1996, as much as one year into the demonstration for some agencies. Second, agencies varied in their degree of compliance in submitting these data to the QA contractor. Chen and Noveck (1998) provide additional discussion of the QA procedures and data.

[10] Because the planned treatments may be partly determined by whether an agency is in the treatment or control group, indicators of planned treatments may be endogenous to patient outcomes and, thus, were not included among the set of additional control variables in the "QA-enhanced" regression models.

TABLE II.3

ADDITIONAL CONTROL VARIABLES AVAILABLE FROM QUALITY ASSURANCE DATA

Comorbid Conditions at Admission

Has Any High Risk Factors
Has Urinary Tract Infection
Has Urinary Incontinence

Severity of Illness at Home Health Admission

Medical Regimen Has Changed Within Past 14 Days
Is Unstable

Functional and Sensory Limitations at Admission

Requires Assistance in Ambulation/Walking
Has Dyspnea
Is Often Confused
Has Depressive Feelings
Displays Depressive Behaviors
Demonstrates Disruptive Behaviors
Is Unable to Prepare Light Meals
Is Unable to Perform Light Housekeeping

Availability of Informal Care at Admission

Assisting Persons Available
Likely that Treatment Can Be Taken Over

Prognosis at Admission

Prognosis Is Good/Fair
Life Expectancy Less than Six Months
Rehabilitative Prognosis Is Good

Treatment Needs at Admission

Requires Assistance with Oral Medications
Surgical Wound Not Healing

form. They identify whether the patient has diabetes, pressure ulcers (stage 3 or stage 4), or a need for complex wound care.

B. FORMING THE ANALYSIS SAMPLE

As noted, the main file contains a number of records that could not (or should not) be part of the sample used to analyze demonstration impacts on home health use. The records excluded from the analysis sample, summarized in Table II.4, fall into two general groups. The first group includes *all* admissions to 12 demonstration agencies that left the demonstration near its start or that had a significant number of unsubmitted or missing claims. The second group of excluded records pertains to patients who were HMO members or had Medicare as a secondary payer on or after their admission date, had missing information on home health use or key control variables, or already had their service use accounted for in prior records. Below, we discuss these exclusions in detail and summarize the resulting analysis sample.

1. Excluding Agencies

We dropped all records from the analysis sample for two control agencies and one treatment agency that left the demonstration near its start. These agencies either were purchased by or merged with another agency, and the new ownership did not want to be part of the demonstration. By dropping these three agencies, we excluded 2,293 records from the analysis sample.

Nine additional agencies, eight treatment agencies and one control agency, had to be dropped from the analysis because they were associated with significant delays in billing during the first demonstration year. Including the agencies could have led to overestimation of reductions in services arising from prospective payment. The process used to identify the agencies, and the

TABLE II.4

PATIENT RECORDS EXCLUDED FROM ANALYSIS SAMPLE, BY TREATMENT STATUS

	Agencies			Admissions			
	Treatment	Control	Percentage of Original Total	Treatment	Control	Total	Percentage of Original Total
Main File[a]	48	43	100.0	61,205	58,153	119,358	100.0
Agency-Level Exclusions							
Participated in Demonstration for Less than One Year	1	2	3.3	348	1,945	2,293	1.9
Missing Bill Data	8	1	9.9	10,219	1,327	11,546	9.7
Patient-Level Exclusions							
Enrolled in HMO	n.a.	n.a.	n.a.	2,276	2,879	5,155	4.3
Has Medicare as Second Payer	n.a.	n.a.	n.a.	1,346	1,537	2,883	2.4
Missing Data	n.a.	n.a.	n.a.	637	1,684	2,321	2.0
Prior Demonstration Admission	n.a.	n.a.	n.a.	4,337	4,375	8,712	7.3
Sample Available for Analysis	**39**	**40**	**86.8**	**42,042**	**44,406**	**86,448**	**72.4**

[a] All demonstration admissions through August 31,1997.

HMO = health maintenance organization; n.a. = not applicable.

sensitivity of our estimates to their exclusion, are discussed in detail in Appendix B, Data Quality. By dropping these nine agencies, we excluded a total 11,546 records from the analysis sample.

After these agency-level exclusions, we were left with 79 agencies, 39 in the treatment group and 40 in the control group. The 79 agencies had 105,519 admissions between their demonstration start date and August 31, 1997, or 88 percent of the total for the 91 originally randomized agencies.

2. Excluding Individual Admissions

We dropped a number of individual patient records. We dropped 5,155 records because HCFA's enrollment files showed that the patients had been enrolled in Medicare HMOs on or after the episode start date. It would have been inappropriate to include these records because the services provided to managed care patients are not subject to prospective payment (or to traditional cost-reimbursement). We dropped another 2,883 records because Medicare was a secondary payer for the patients on or after the episode start date. As in the case of managed care patients, the services provided to these patients are not subject to per-episode payment, so that agencies' behavior should be largely independent of the demonstration incentives.

We dropped another 2,321 records because data were missing. In the case of approximately half the records, the data fields of their UB-92 forms did not contain complete information on patient characteristics. This information had been inadvertently erased from the forms near the start of the demonstration because of an error in the software used by PGBA. Although we used data from the QA admission forms in an attempt to replace the missing information, only a small number of records could be repaired, because the QA forms were not collected until May 1996. The remaining records from the group of 2,321 with "bad" data did not show any home health services during the first month after admission. In most cases, these records reflected "false" admissions, in which the only service provided was durable medical equipment.

We also excluded records from the analysis sample if our file already contained valid patient records from a previous episode start. In total, 8,712 records for 8,146 patients were dropped for this reason. As discussed in Chapter III, these subsequent admissions were dropped because they could have led to incorrect conclusions about the impacts of prospective payment on service use.[11] Because each record captures service use from the point of admission through the end of our analysis period, all the care identified on any later records was already accounted for on the initial record for the patient.

3. Analysis Samples

After dropping all these records from the main file, we arrived at an analysis sample that contained 86,448 unique patient admissions to 79 demonstration agencies (40 in the control group and 39 in the treatment group). This sample reflects roughly 72 percent of all the admissions taking place in demonstration agencies during our analysis period.

As shown in Table II.5, a majority of the records in the analysis sample--52,133--pertain to admissions during an agency's first year in the demonstration. These records are about evenly split between those in treatment agencies (26,384) and in those control agencies (25,749). They contain at least 12 months of information on service use subsequent to the episode start; most have 16 months of information.[12] Nearly all the remaining records--32,198--reflect admissions in the second demonstration year. Most of these year-two records provide eight months of data on postadmission

---

[11] The basic concern is that prospective payment provides an incentive for treatment agencies to distribute care over multiple episodes and obtain multiple episodic payments from the same patient. This incentive could lead to bias in our impact estimates unless we were to count service use over a fixed period from the start of the *initial* admission.

[12] A larger proportion of the admissions data of control agencies cover 16 months because more of the early (pre-January 1996) entrants in the demonstration were in the control group.

TABLE II.5

NUMBER OF PATIENT RECORDS IN ANALYSIS SAMPLE THAT COVER SELECTED TIME PERIODS FROM ADMISSION, BY DEMONSTRATION YEAR AND TREATMENT STATUS

	Total (Available) Demonstration Period		Year One		Year Two		Year Three	
Number of Records with Complete Data from Episode Start (Admission) Through:	Control Group Patients	Treatment Group Patients	Control Group Patients	Treatment Group Patients	Control Group Patients	Treatment Group Patients	Control Group Patients	Treatment Group Patients
4 Months (120 Days)	44,406 (40)	42,042 (39)	25,749 (40)	26,384 (39)	17,553 (40)	15,365 (39)	1,104 (12)	293 (5)
8 Months (240 Days)	38,137 (40)	35,940 (39)	25,749 (40)	26,384 (39)	12,388 (40)	9,556 (39)	--	--
1 Year (365 Days)	30,576 (40)	28,062 (39)	25,749 (40)	26,384 (39)	4,827 (36)	1,678 (33)	--	--
16 Months (480 Days)	23,913 (40)	20,870 (39)	22,547 (40)	20,508 (40)	1,366 (13)	362 (7)	--	--

SOURCE: Medicare claims data.

NOTE: Numbers in parentheses are the number of agencies.

use, but only 6,505 records contain 12 months of data, and only 1,728 contain 16 months.[13] The analysis sample contains only a small number of records for demonstration year three. These records have information for only the first 120 days of care (the at-risk period), and they include just 17 demonstration agencies.

Because the patient records provide varying amounts of postadmission data, we typically must restrict the sample that we use for a given analysis. (Chapter III fully discusses this issue.) For example, in much of our main impact analysis presented in Chapter IV and in our subgroup analysis presented in Chapter V, we examine home health use over a period of 12 months from admission. As shown in Table II.5, we have 12 months of complete data for all agencies for only the first demonstration year. Thus, given this data constraint, we restrict the sample for these analyses to admissions in that demonstration year.

For each agency, the number of patient admissions in our analysis sample ranges quite dramatically. For the first year of the demonstration, for example, the largest treatment agency had roughly 4,500 patient admissions in our sample, and the largest control agency had about 3,500 admissions (Table II.6). These two agencies account for about 17 percent and 14 percent of the entire samples in their respective groups. In contrast, all six of the smallest agencies (three in each group) had between 30 and 52 admissions, each no more than 0.2 percent of their respective samples. On average, the number of admissions by treatment agencies (677) was slightly larger than average number by control agencies (644). If one ignores the largest agency in each group, however, the

---

[13] The number of agencies shown in Table II.5 as providing a full year of data for demonstration year two is somewhat misleading because the list includes agencies that entered the demonstration on January 1, 1996, and that had an admission on the same day. When this group is excluded, the number of year-two admissions drops by less than 200 but the number of agencies drops from 69 to only 34.

TABLE II.6

DISTRIBUTION OF ADMISSIONS IN ANALYSIS SAMPLE FOR THE FIRST DEMONSTRATION YEAR, BY TREATMENT STATUS

Admissions by an Agency	Treatment Group	Percentage of All Treatment Admissions	Control Group	Percentage of All Control Admissions
Most	4,497	17.0	3,508	13.6
Second Most	2,440	9.2	1,711	6.6
Third Most	1,807	6.8	1,594	6.2
Third Fewest	49	0.2	52	0.2
Second Fewest	37	0.1	46	0.2
Fewest	30	0.1	46	0.2
Total Admissions	**26,384**	**100.0**	**25,749**	**100.0**
Agency Average	**677**	**2.6**	**644**	**2.5**

remaining agencies had nearly identical admissions by (576 for treatment agencies and 570 for control agencies; data not shown).

C. SUMMARY STATISTICS FOR THE MAIN CONTROL VARIABLES

We used regression models to estimate demonstration impacts largely because these models control for any preexisting differences between treatment and control agencies that might have otherwise biased our results. These differences could have existed despite random assignment because the number of agencies is fairly small and several agencies had to be dropped from our analysis. (Details on these regression models, and related methodological issues, are addressed in Chapter III.)

1. Characteristics of the Sample

The control variables for our main regression models, listed in Table II.2, include an array of key patient- and agency-level characteristics. In Table II.7, we display the sample means of these variables for patients admitted during the first demonstration year.[14] The means are presented for the entire sample, and also separately for the treatment and control groups in order to examine whether the two groups differed in any important ways.[15,16]

---

[14] As noted, data limitations required us to restrict the sample for much of our analysis to patients admitted during the first demonstration year.

[15] The means were constructed using sample weights that gave each agency equal importance in the calculations, as this approach was the one used to estimate impacts. The significance levels for the tests of equality between the treatment and control group means did not account for design effects due to clustering because we were interested only in differences within this sample, not the population of all agencies. (This is not the case in the main analysis.) We did account for the design effects associated with our use of sample weights, however. Chapter III fully discusses the use of weighting and clustering in our analysis.

[16] With the large sample available, sufficient statistical power exists to detect very small differences between the treatment and control groups at baseline. We therefore expected that many differences in the explanatory variables between the treatment and control agencies would be statistically significant, even when the magnitude of the difference was not materially important.

TABLE II.7

WEIGHTED MEANS FOR EXPLANATORY VARIABLES, BY TREATMENT STATUS, AND TESTS FOR DIFFERENCES IN TREATMENT AND CONTROL GROUP MEANS

Explanatory Variable	Total	Treatment Group	Control Group
Demographic Measures (Percentage)			
Age Was Original Reason for Medicare Eligibility	82.5	83.9	81.2***
Age: <65 Years	8.3	7.5	9.0***
Age: 65 to 74 Years	29.3	28.5	30.1**
Age: 75 to 84 Years	39.9	40.2	39.6
Age: >84 Years	22.6	23.9	21.3***
Race/Ethnicity: White	81.5	81.7	81.2
Female	64.4	64.7	64.1
Enrolled in Medicaid	25.3	26.4	24.1***
Medical Conditions and Care (Percentage)			
Cancer	13.1	13.2	13.0
Diabetes	21.4	21.3	21.5
Cerebrovascular Accident (Stroke)	14.5	14.6	14.5
Decubitus Ulcer: Stage 3 or 4	4.2	4.6	3.8**
Needs Complex Wound Care	6.7	6.7	6.6
Any Medical Condition	46.1	46.4	45.8
Limitations in ADLs (Percentage)			
Bathing	72.8	72.9	72.8
Eating	28.7	28.0	29.3*
Dressing	62.8	61.4	64.2***
Toileting	39.1	38.0	40.3***
Transferring	51.0	50.4	51.6
Any ADL	78.0	78.2	77.8
Patient Prior Service Use Measures			
Admitted to Home Health from Hospital (Percentage)	36.4	34.1	38.7***
Length of Hospital Stay Ending Within 14 Days Before Home Health Admission (Number of Days)	3.9	3.6	4.2***
Had SNF Stay Ending Within 14 Days Before Admission (Percentage)	16.7	17.9	15.4***
Total Part A (Plus Part B Inpatient/SNF/Home Health/Hospice) Reimbursement in Six Months Before Admission (Dollars)	11,222	11,278	11,167
Part A Home Health Visits in Six Months Before Admission (Number)	12.2	11.9	12.5

TABLE II.7 *(continued)*

Explanatory Variable	Total	Treatment Group	Control Group
Other Patient-Level Variables (Percentage)			
Had Medicare for Less than Six Months Before Admission	1.4	1.4	1.4
Enrolled in HMO in Six Months Before Admission	0.9	0.9	0.9
Had Medicare as Secondary Payer in Six Months Before Admission	0.3	0.2	0.3
Admission Date: 5/30/95 Through 9/30/95	5.6	4.1	7.1***
Admission Date: 10/1/95 Through 12/31/95	10.2	8.9	11.5***
Admission Date: 1/1/96 Through 3/31/96	28.2	27.4	29.0**
Admission Date: 4/1/96 Through 6/30/96	24.7	26.0	23.4***
Admission Date: 7/1/96 Through 9/30/96	18.2	19.0	17.4**
Admission Date: 10/1/96 Through 12/31/96	13.1	14.6	11.6***
Agency Characteristics Measures (Percentage)			
Proprietary	50.6	48.7	52.5***
Hospital Based	11.4	7.7	15.0***
Chain	35.4	43.6	27.5***
Small Agency (<30,000 Visits in Base Year)	27.8	35.9	20.0***
Agency Practice Pattern: Index of Average Visits[a]	1.03	0.95	1.11***
Area Characteristics Measures			
California (Percentage)	22.8	23.1	22.5
Florida (Percentage)	7.6	7.7	7.5
Illinois (Percentage)	19.0	15.4	22.5***
Massachusetts (Percentage)	12.7	17.9	7.5***
Texas (Percentage)	38.0	35.9	40.0***
Urban (Percentage)	83.5	82.1	85.0***
Physicians per 10,000 Persons (Number; 1994)	21.5	21.6	21.5
Nursing Home Beds per 100 Persons >65 Years (Number; 1991)	5.2	5.2	5.2
Hospital Occupancy Rate (Percentage; 1993)	61.5	62.2	60.8***
Sample Size			
Patients	**52,133**	**26,384**	**25,749**
Agencies	**79**	**39**	**40**

[a] An index of the case-mix-adjusted average visits received by an agency's patients during the first 120 days of base-quarter episodes, relative to the average across all agencies.

*Difference in means are significantly different from zero at the .10 level, two-tailed test.
**Difference in means are significantly different from zero at the .05 level, two-tailed test.
***Difference in means are significantly different from zero at the .01 level, two-tailed test.

a. Demographic Measures

More than 80 percent of the patients in the sample were eligible for the Medicare program because of their age, and more than 90 percent were older than age 65. The largest group, roughly 40 percent, was between the ages of 75 and 84. About 82 percent of the sample was white, 64 percent female, and 25 percent jointly enrolled in the Medicaid program Some of the differences in demographic characteristics between treatment and control agencies were significant, but the magnitudes of these differences was fairly small. Relative to control agencies, treatment agencies served a somewhat higher proportion of patients older than age 85, enrolled in Medicaid through a state buy-in program, or who became eligible for Medicare because of their age.

b. Medical Conditions and Care

At admission, about one in seven (13 percent) of the patients in the sample had cancer, and a slightly higher percentage (15 percent) was recovering from a stroke. About one in five patients (21 percent) had diabetes. A very small percentage of patients had a serious decubitus ulcer (four percent) or required care for a complex wound (seven percent). The rates of these conditions were very similar for patients in the treatment and control agencies. Although patients in treatment agencies had a significantly higher rate of decubitus ulcer, the difference was less than one percentage point.

c. Limitations in Activities of Daily Living

A large majority of patients had one or more limitations in ADLs for which we have measures. The most common limitations were in bathing (73 percent) and dressing (63 percent); the least common was in eating (29 percent). Patients in treatment agencies had significantly lower rates of limitation for three of the five ADL tasks for which we have measures (dressing, toileting, and

transferring). However, all three differences were small, from about four to six percent of the control group mean.

d. Patient Prior Service Use Measures

More than one-third (36 percent) of patients had been hospitalized during the two weeks preceding admission, and another 17 percent had been in a skilled nursing facility (SNF). Means for both characteristics differed significantly between treatment and control agencies; treatment agencies had a lower share of patients from the hospital, but a larger share of patients leaving an SNF. Both these differences are explained mainly by the larger number of hospital-based agencies in the control group than in the treatment group (discussed in the following section). Not surprisingly, this subgroup admitted a disproportionate number of its patients from the hospital.

e. Agency Characteristics Measures

Roughly half the patients in the total (weighted) sample were admitted by proprietary agencies, while only 11 percent were admitted by hospital-based agencies. About one-third of patients in the total sample had been admitted by agencies that were part of a chain, and roughly one-quarter (28 percent) were in small agencies, defined as providing fewer than 30,000 visits in the base year prior to the demonstration. Finally, patients in the sample were admitted by agencies with an average practice pattern indexed at 1.03, which is just slightly above the mean of 1.00 for all agencies in the demonstration.[17]

---

[17] Recall that the practice pattern variable accounts for preexisting differences in the number of visits provided to patients, controlling for variation in case-mix distribution. An agency with an index value above 1.0 is defined as having a high-use practice pattern since, in comparison to other agencies in the demonstration, it provided more visits to a typical patient in the base year. Similarly, an agency with an index below 1.0 is defined as having a low-use practice pattern. The average index value of 1.03 (rather than 1.0) shown in Table II.7 results from the exclusion of a small number of demonstration agencies from the sample (see Appendix B for details).

Significant treatment-control differences existed for each of these agency characteristics. A significantly smaller proportion of treatment group patients had been admitted by proprietary or hospital-based agencies, while a significantly higher proportion of these patients had been admitted from a small agency or a chain agency. In addition, the average patient in the treatment group had been admitted by an agency with a practice pattern about 14 percent below the average patient in the control group (0.95 compared to 1.11). This difference is quite striking, since it essentially indicates that the average treatment patient was in an agency that provided about 14 percent fewer visits to a typical patient before the demonstration even began.

f. Area Characteristics Measures

The largest share of patients in the overall sample (38 percent) were from Texas, followed by California (23 percent) and Illinois (19 percent), Massachusetts (13 percent), and Florida (8 percent). More than 80 percent of patients were admitted to agencies located in urban areas. Significant treatment-control differences were evident in these area characteristics. In particular, treatment group patients were significantly more likely to be admitted from agencies in Massachusetts, while they were significantly less likely to be admitted from agencies in Illinois and Texas.

2. Implications for the Analysis of Program Impacts

Although the preexisting treatment-control differences in patient characteristics are minimal, there are several large differences in agency and area characteristics.[18] These differences underscore the importance of using regression models to estimate program impacts, as these models control for effects on service use that might otherwise be falsely attributed to prospective payment. For

---

[18] Some agency-level differences between treatment and control groups are partly due to the removal of 12 of the demonstration agencies from the analysis sample (discussed previously). However, the large differences in practice patterns and state of operation would have remained even with all 91 agencies in the sample.

example, the far lower service use by a typical patient in a agency prior to the demonstration (as measured by the index of practice patterns) suggests that a simple comparison of mean visits under the demonstration might substantially overstate the program impact. In the next chapter, we describe the regression models used to isolate the impact of the method of payment and address other issues involved in accurately estimating program effects.

III. METHODS

Our analysis for this report focuses on how prospective per-episode payment affects the provision of home health services relative to a payment system of cost-reimbursement. To conduct this analysis, we must address two key methodological issues. First, we must properly define the units of measurement for each outcome; second, we must use appropriate statistical techniques to estimate demonstration impacts.

A. UNITS OF ANALYSIS

Recall that a complete patient episode under the demonstration includes an initial 120-day ("at-risk") period for which agencies receive a single prospective (episodic) payment, followed by an "outlier" period for which they receive prospective per-visit payments. This episode forms a natural unit of analysis as it presumably encompasses all the care an agency provides for a given spell of illness. However, as discussed in this section, an impact analysis based on the patient episode could lead to misleading estimates because an agency's treatment status may affect the number of episodes over which care is distributed. To avoid this complication, we will generally examine demonstration impacts over fixed periods occurring after the initial patient admission. These periods, summarized for key outcomes in Table III.1, enable us to estimate unbiased measures of program impacts because they are unaffected by the demonstration incentives.

1. Problem with Investigating Impacts Within the Patient Episode

If treatment agencies respond to the demonstration by reducing episode length, using an episode of care as the unit of analysis could overstate demonstration impacts. Consider, for example, a patient who was admitted to a control agency, received one visit every other day for 300 days (150

TABLE III.1

UNITS OF MEASUREMENT FOR KEY OUTCOME VARIABLES IN THE IMPACT ANALYSIS

Key Outcome Variables	Unit of Measurement			
	Per Month or Per Semi-Month	Through 120 Days After Admission (At-Risk Period)	In Months 5-8 and 9-12 After Admission	Within One Year After Admission
Service Use				
Number of Visits, Total	x	x	x	x
Number of Visits, by Type	--	x	x	x
Probability of a Visit, Any	x	--	x	--
Probability of a Visit, by Type	--	x	x	x
Service Duration				
Days Until Last Visit	--	--	--	x
Probability of a Visit After 120 Days	--	--	--	x

visits total), and was discharged. If this patient were admitted to a treatment agency, the incentives of the demonstration might lead that agency to provide the same services (that is, one visit every other day) but to discharge the patient earlier. Consider, therefore, a similar patient who was admitted to a treatment agency, received one visit every other day, but was discharged after only 100 days. Because of this earlier discharge, there may be a greater chance that the patient will have to be readmitted, leading to a new episode start.[1] Assume that the readmission occurred on day 201 (from the original episode start), and that the treatment agency again provided one visit every other day through day 300, after which it discharged the patient. Therefore, the control agency provided a total of 150 visits during one episode that lasted 300 days, whereas the treatment agency provided two 50-visit episodes that each lasted 100 days.

If we use the episode as the unit of analysis in this example, then a simple treatment-control difference would lead us to conclude that the demonstration reduced the number of visits from 150 to 50, or by 67 percent. However, the *actual reduction in use* brought about by the demonstration was only 150 to 100 visits, or 33 percent. By measuring use within an episode of care, we have overstated the demonstration impact because the treatment agency has "spread" its care over multiple episodes.

During the first year of the demonstration, treatment agencies did have slightly higher rates of multiple episodes for the same patient. Within a 16-month period after the first episode start, the average rate of second episodes was 10.2 percent in control agencies and 10.9 percent in treatment agencies. This difference was not statistically significant, but it could nevertheless lead to a small

---

[1] The need for readmission by the treatment agency does not necessarily imply that inadequate care has been provided. The readmission might result from a second "spell" of health problems, for example, or from complications that also would have arisen had the patient been admitted to the control agency. The difference is that the control agency had not discharged the patient, so the services for this second spell would be counted within the original episode.

overstatement of demonstration impacts. It thus underscores the need to focus on service use measures that are independent of the system of payment.

2. Main Unit of Analysis: The "Patient Year"

To avoid the problem associated with using episodes of care as the unit of analysis, we focus on patients' service use during a fixed period after admission (the episode start). In principle, this period could be of any duration; however, we face a tradeoff between increasing duration and losing observations because of incomplete data. After taking this tradeoff into account, we have settled on using a "patient year"--the first 365 days from the episode start--as the period for our main analysis.

To understand more completely how impacts may be taking place, we also examine impacts within each "trimester" of the patient year. Specifically, we examine impacts during (1) the first four months (120 days) of the patient year--that is, the at-risk period; (2) the subsequent four months of the patient year, from day 121 through day 240; and (3) the final four months of the patient year, from day 241 through day 365.

Depending on the start date of a given agency, we have a full year of data on service use for anywhere from 12 to 19 months of admissions. The benefit of using all these admissions is that it improves the precision of our impact estimates.[2] However, because agencies' behavior may be related to how long they participated in the demonstration, our estimates should be based on a consistent period of participation for all agencies. For our main analysis, we will therefore limit our sample to admissions taking place within the agencies' first 12 months in the demonstration.[3]

---

[2] The gains in precision may actually be quite modest after accounting for the "clustering" of admissions within the demonstration agencies. This issue is addressed in the next section.

[3] Data for admissions in the second demonstration year indicate that both treatment and control agencies experienced decreases in the number of visits as the demonstration continued (see Chapter

(continued...)

Because we are using patients, rather than episodes, as the unit of analysis, we treat subsequent episodes that a given patient experiences as additional utilization and define a patient year *only* from the first admission that we observe for the patient. By doing so, our impact estimates are unaffected by the possible correlation between treatment status and subsequent episode starts, and we avoid double counting services. This approach will lead us to "miss" some amount of care that is subject to prospective payment; however, the amount should be very small, as there are relatively few subsequent episodes, and much of the care in these episodes should be picked up in a patient year.

a. Comparing Demonstration Impacts over Time

Because we are interested in whether demonstration impacts change over time, we will compare impacts during the agencies' first and second years in the demonstration. The start date of each episode is used to determine the demonstration year in which the observation falls. If the start date occurs after the 365th day since an agency entered the demonstration, it falls within the second demonstration year; otherwise it falls within the first.[4] This process leads to the creation of two separate samples, corresponding to demonstration year one and demonstration year two.

For the majority of "late" agencies, which entered the demonstration on January 1, 1996, we had a full 12 months of postadmission data on only a few patients admitted during the second

---

[3](...continued)
VI). Because the sample of "early" agencies (those beginning the demonstration in 1995) includes an unequal number of treatment and control agencies (12 and 21, respectively), we are particularly concerned that including admissions from demonstration year two could have misstated our main impact estimates. In Chapter VI, we present a separate analysis of how impacts changed between the first and second years of the demonstration.

[4] For six agencies that entered the demonstration before August 31, 1995, we have a small number of admissions in demonstration year three for which we observe service use for at least 120 days postadmission. We will not investigate these data because there are too few observations to enable us to draw valid conclusions.

demonstration year, and, in the case of some of those agencies, on no patients (see Figure III.1).[5] The comparison of impacts between demonstration years must therefore be limited to periods of care shorter than a patient year. The two periods that we examine are: (1) the first 120 days from the episode start (the at-risk period); and (2) the first eight months from the episode start. To achieve the highest possible precision in our estimates, each of these analyses uses all of the (initial) admissions for which we have sufficiently complete service data.[6] For the analysis of the at-risk period, we have between 8 and 12 months of admissions for all agencies in the second demonstration year; for the analysis of the first eight months, we have between 4 and 10 months of admissions.

One issue associated with this analytic approach is that the subsample for year-two admissions differs substantively from the subsample for demonstration year one. Specifically, the set of admissions for demonstration year two excludes patients with admissions during the preceding year (in demonstration year one) because those admissions were subject to prospective payment. In sensitivity analyses (not reported), we compared selected impacts for the full sample in demonstration year one with the impacts for a subsample of patients who had no admissions to demonstration agencies in the prior (predemonstration) year. These impacts were quite similar, suggesting that any impact differences between demonstration years one and two would not be the result of sample construction.

---

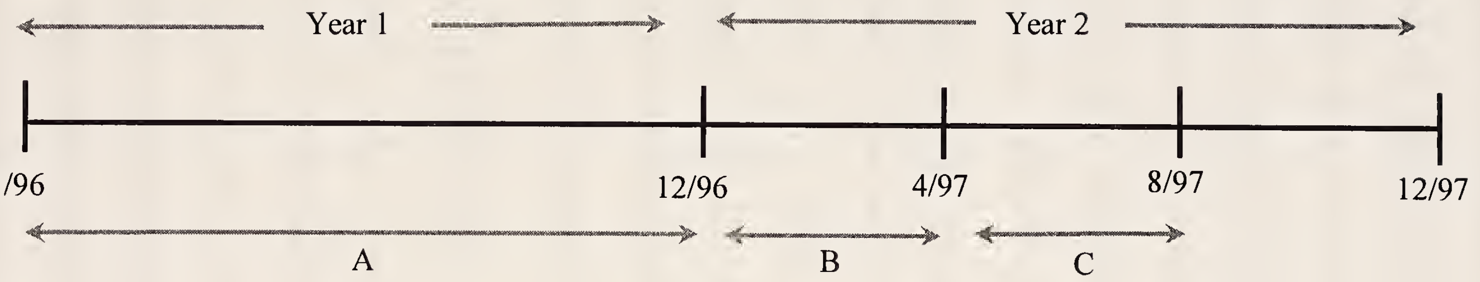
[5] Specifically, we have a complete patient-year of data for 228 patients admitted to 35 (of the 45) late agencies on the agencies' first day in the demonstration (January 1, 1996).

[6] We are less concerned about consistently defining the samples for demonstration year two because our goal is simply to identify changes in impacts between years. To the extent that we find any changes taking place, we can further evaluate how they would differ if we defined the agency samples over equal periods from the start of the demonstration.

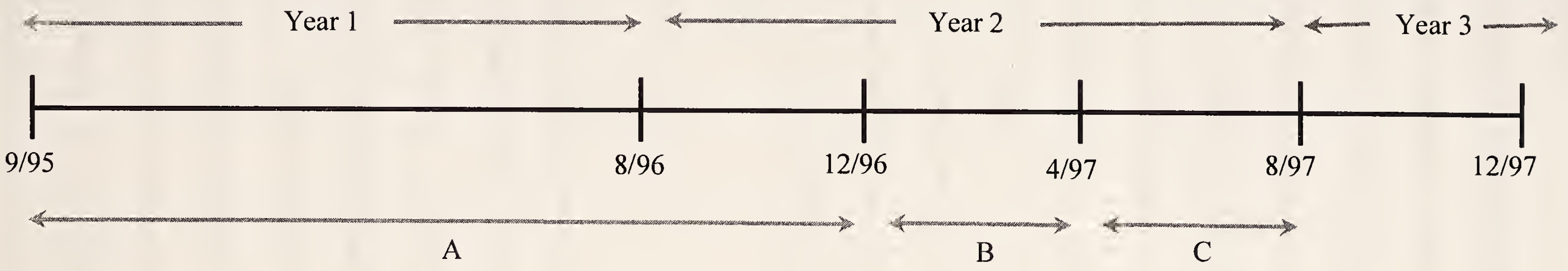
FIGURE III.1

AVAILABILITY OF DATA TO COMPARE IMPACTS BETWEEN DEMONSTRATION YEARS

Agencies Beginning the Demonstration on January 1, 1996



Agencies Beginning the Demonstration "Early," Before January 1, 1996[a]



For admission occurring in interval:

- A = Full 12 months (patient-year) of postadmission data available
- B = Less than 12 months of postadmission data available. All admissions complete through at least 8 months
- C = Less than 8 months of postadmission data available. All admissions complete through at least 4 months

[a] Early agencies began the demonstration between June 1, 1995, and November 1, 1995. The agency start date shown (September 1, 1995) is for illustrative purposes.

b. Investigating Service Use Within the Episode of Care

The episode of care, although not the focus of our impact analysis, still provides several valuable pieces of information. First, by examining the impact of the demonstration on the number of episodes per patient, the rate of episodes with only one or two visits, and the number of outlier visits, we gain a more detailed understanding of how treatment agencies responded to prospective payment. This understanding is important because any cost savings from lower service use under prospective payment may be offset by increases in admissions or readmissions, or by an expansion in services provided during the outlier period. Second, by analyzing detailed descriptive data on the length of episodes, the number of at-risk and outlier visits per episode, and other outcomes, we gain an understanding of how care is distributed under a system of prospective payment.

Because this analysis of the patient episode is largely concerned with describing patterns of care, we examine use over a 16-month period from admission. This longer period enables us to observe the at-risk period and a full 12 months of the outlier period.

B. STATISTICAL TECHNIQUES

We use regression models to estimate demonstration impacts because they control for preexisting differences between treatment and control agencies that may exist despite random assignment, and because they improve statistical precision. For our main findings, we estimate impacts using ordinary least squares (OLS) regression when the dependent variable is continuous and logistic regression when the dependent variable is binary. As part of the sensitivity analysis, we also estimate impacts using Tobit regression models when the dependent variable is "censored," that is, has a natural (or imposed) lower or upper bound at which a portion of the sample is clustered. (The number of physical therapy visits, for example, is censored at zero.) The main regression models use weighted data that give each agency equal representation in the analysis; however, as a

sensitivity test, we also examine impacts in the absence of sample weights. All analyses use standard errors that take into account the effects of sample clustering and weighting, as described in Section B.5 and B.6.

1. Regression Models for Estimating Overall Impacts

When the outcome that we investigate is continuous (for example, number of visits or duration of services), the basic model that we use to estimate the overall impacts of prospective payment is:

$$(1) \quad Y = \alpha + X\beta + \delta T + \epsilon,$$

where:

Y = a continuous outcome variable, measured after a patient's admission to a demonstration home health agency

X = a matrix of control variables

T = a binary variable for treatment status that equals one for admissions to treatment agencies and zero for admissions to control agencies[7]

α = the intercept term

β = the vector of regression coefficients on the control variables

δ = a regression coefficient on the variable for treatment status

ϵ = a random disturbance term assumed to have a mean of zero (conditional on the regressors in the model) that reflects all the unobserved factors affecting the outcome variable

---

[7] The omitted (reference) binary variable equals one for admissions to control agencies and zero for admissions to treatment agencies.

The coefficient δ on the variable T measures the effect of the demonstration payment method on the (continuous) outcome of interest and is tested to determine whether it is significantly different from zero. For the main analysis, we use OLS to estimate this equation; however, when the dependent variable is censored, we also use Tobit models to investigate impacts (see Section B.7.c).

When the outcome variable that we investigate is binary (for example, whether a patient received any care or whether an episode extends into the outlier period), the logit model is used to estimate demonstration impacts. The structure of the logit model is as follows:

$$(2) \quad Probability\ (Y = 1) = \frac{1}{1 + e^{(\alpha + X\beta + \delta T)}},$$

where Y is the (binary) outcome variable and the remaining variables and parameters are defined as in equation (1).[8]

Given the nonlinearity of the logit model, the estimated impact of the payment method is not measured directly by the coefficient δ on the variable for treatment status. Instead, to estimate the demonstration impact on the probability that $Y = 1$, we use the coefficient estimates from the model to generate two predicted probabilities for each observation: one assuming that the observation belongs to the treatment group ($T = 1$), and one assuming that it belongs to the control group ($T = 0$). The impact estimate is the average difference between these estimated probabilities. Because the statistical significance of δ determines whether the odds that $Y = 1$ are significantly different for

---

[8] For ease of exposition, all later specifications in which the outcome variable may be continuous or discrete are shown for the continuous outcomes only. In all cases, however, we use both OLS and logit models as appropriate.

the treatment and control groups, we use the p-value on this parameter to test our hypotheses about differences between the two groups.

Throughout the tables of results, we present the means of each outcome variable for the treatment group and the control group alongside the estimated impact, as points of reference. The control group mean is unadjusted and reflects a reasonable estimate of the mean value expected to occur in the absence of the demonstration. The treatment group mean is regression adjusted by setting it equal to this control group mean minus the (regression-based) impact estimate.

2. Estimating Subgroup Impacts

We investigate demonstration impacts for subgroups for two reasons. First, if impacts for the full sample exist, we are interested in determining whether they differ between key subgroups of agencies or patients. Second, even if we find no compelling evidence of overall impacts, it is important to determine whether they exist for selected subgroups.

The subgroup analysis uses a regression model similar to equation (1). The only difference is that the subgroup model includes additional control variables formed by interacting treatment status T with a set of binary (subgroup) variables defined from agency or patient characteristics. For the agency subgroup analysis, we estimate a single equation with a set of five subgroup variables interacted with treatment status. Each subgroup variable included in the regression reflects a key agency characteristic (for example, agency is for profit), and its corresponding (reference) group reflects the absence of this characteristic (for example, agency is nonprofit).[9] For the patient-

---

[9] As described in Chapter V, the subgroup "pairs" that we examine are as follows: (1) for-profit status or nonprofit status; (2) high-use practice patterns vs low-use practice patterns; (3) small size or large size; (4) freestanding or hospital based; and (5) above or below the base-year cost limits.

subgroup analysis, we estimate three regression equations, with only one patient subgroup pair being examined at a time.[10]

The basic regression model for subgroups is given by:

$$(3)\quad Y = \alpha + X_1\beta_1 + X_2\beta_2 + \delta T + \sum_{i=1}^{n} \gamma_i \, (X_{1i} * T) + \epsilon,$$

where:

Y = an outcome variable

X_1 = a matrix of n subgroup variables, each of which takes on a value of one if the observation belongs to the subgroup and zero otherwise[11]

X_2 = a matrix containing all other control variables

T = a binary variable for treatment status that equals one for admissions to treatment agencies and zero for admissions to control agencies

β_1 and β_2 = vectors of regression coefficients on the variables in X_1 and X_2, respectively

γ_i = a regression coefficient on the interaction term between X_{1i} and T[12]

α = the intercept term

---

[10] We evaluate three pairs of patient-level subgroups: (1) patient can take oral medications independently or patient cannot take oral medications independently; (2) patient has other caregivers available or patient does not have other caregivers available; and (3) patient has high expected costs or patient has low expected costs. Inability to take oral medications is a very strong predictor of high service use, which may lead any impact differences that exist within the oral medications subgroup to be distorted under a regression model that estimates patient subgroup impacts simultaneously. For this reason, we choose instead to estimate a separate regression model for each of the three patient subgroup pairs. A similar issue may arise in the agency subgroup analysis; for example, we observe that for-profit status is strongly correlated with high-use practice patterns. However, in this case, it is desireable to control for such correlation in order to isolate which (if any) features lead to significant differences in agencies' response to demonstration incentives.

[11] For the agency subgroup model, n equals five; for each (separate) patient subgroup model, n equals one.

[12] This interaction term takes on a value of one if the observation belongs to the ith subgroup *and* is a treatment admission; and zero otherwise.

δ = a regression coefficient on the variable for treatment status

ϵ = a random disturbance term assumed to have a mean of zero (conditional on the regressors in the model) that reflects all the unobserved factors affecting the outcome variable

To assess whether the impacts between particular subgroups differ, we examine the statistical significance of the coefficient γ_i on the corresponding interaction term. For example, suppose that the *p*-value for the coefficient on the interaction term between an agency's for-profit status and its treatment status is .04 (statistically significant from zero). This test indicates that the effect of prospective payment on that outcome measure differs significantly between for-profit and nonprofit agencies.

We also estimate the effect of prospective payment for each individual subgroup.[13] The impact for a subgroup is estimated by setting the indicator variable for the subgroup appropriately and then obtaining the predicted value of the outcome variable for each observation, first as if it were from a treatment agency ($T = 1$), and then as if from a control agency ($T = 0$). The mean difference between these two predicted values provides an unbiased estimate of the impact of prospective payment on the outcome Y for each subgroup. Thus, for the subgroup of agencies/patients with characteristic j, the impact estimate is given by[14]:

$$(4) \quad \delta + \sum_{i \neq j} \gamma_i \overline{X_{li}} + \gamma_j X_j,$$

---

[13] It may not be possible to identify small or moderate impacts for subgroups, as tests of statistical significance lose power as the sample size decreases. The smaller the size of the subgroup (all else equal), the less likely we are to reject the hypothesis that the payment method has no effect, for any given true effect size.

[14] For the patient-level subgroup analysis, the expression simplifies to: $\delta + X_j \gamma_j$.

where $\overline{X}_{Ii}$ equals the sample mean of the ith subgroup variable X_{Ii}. A hypothesis test is conducted on the linear combination given by equation (4) to determine the significance level of the estimated demonstration impact for each subgroup. Results from each test appear in the tables of Chapter V, as the *p*-value for the treatment-control difference within each subgroup. (For more information on hypothesis testing, see Section 4.)

3. Investigating Differences in Impacts Between Demonstration Years

To examine whether the impact of the demonstration changes with agencies' length of participation, we combine the available sample from demonstration year two with our main analysis sample from demonstration year one and estimate a single regression model for each outcome of interest. The specification of this model is similar to those shown in equations (1) and (2) for continuous and dichotomous outcomes, respectively; the only difference is that the specification includes an additional dummy variable for the demonstration year in which the admission took place, and an interaction term between the agency's treatment status and the demonstration year. That is:

$$(5) \quad Y = \alpha + X\beta + \delta T + \lambda_1 D + \lambda_2 (T * D) + \epsilon,$$

where:

Y = an outcome variable

D = a dummy variable that equals one if the admission took place in demonstration year two, and zero if it is from demonstration year one

X = a matrix of control variables

T = a binary variable for treatment status that equals one for admissions to treatment agencies and zero for admissions to control agencies

λ_1 = the coefficient on the dummy variable D

λ_2 = the coefficient on the interaction term between the variables D and T, where this interaction term equals one for observations reflecting admissions in the second demonstration year to treatment agencies, and zero otherwise

α = the intercept term

β = the vector of regression coefficients on the control variables

δ = a regression coefficient on the variable for treatment status

ϵ = a random disturbance term assumed to have a mean of zero (conditional on the regressors in the model) that reflects all the unobserved factors affecting the outcome variable

As before, the estimated impact of prospective payment in the first demonstration year is given by the coefficient δ. The estimated impact in second demonstration year is given by the sum of coefficients δ and λ_2. The difference between these two impacts, λ_2, then measures the change in the effect of prospective payment from the first to the second year of the demonstration.

4. Hypothesis Tests for the Impact Estimates

For each outcome, a two-tailed t-statistic tests the null hypothesis that there is no difference between the regression-adjusted population means for treatment and control agencies.[15] The associated p-value, which indicates the probability of obtaining a sample estimate of the observed magnitude if the null hypothesis were true, is used to determine whether the demonstration had a measurable impact. The p-value is based on estimated standard errors that account for the clustering of patients within agencies and the use of sample weights. A p-value of less than .10 indicates rejection of the null hypothesis and provides significant statistical evidence that a demonstration

---

[15] Two-tailed tests are used throughout our analysis to avoid confusion and to flag estimates of the "wrong" expected sign that are large enough to be statistically significant. For impacts with the "correct" expected sign, a two-tailed test is less likely than a one-tailed test to reject the hypothesis of no demonstration effect (all else equal).

impact probably exists. At this *p*-value, however, approximately 10 percent of independent tests will show, simply by chance, a statistically significant treatment-control difference when there is no true program effect (known as a Type I error). Therefore, in assessing whether a statistically significant treatment-control difference, especially one with a *p*-value between .05 and .10, should be interpreted as a true program impact, we consider whether the sign and magnitude of the predicted effect are consistent with those for related outcomes.

Despite our large sample of patient admissions, it is unlikely that small treatment-control differences will be statistically significant because design effects (discussed in Section B.5) greatly reduce the precision of our estimates. For example, if we ignore the design effects associated with our weighted sample and the clustering of patients within agencies, we have an 80 percent probability of detecting an impact on total visits per patient of less than one percent under a two-tailed test at the 10 percent significance level. After accounting for design effects, however, this minimum detectable effect is approximately eight percent. Thus, despite our large sample size, it is unlikely that we would detect the effects of prospective payment on overall visits, or other outcomes, unless they were at least moderate in size. Nonetheless, in most cases, the sample is adequate to detect impacts of policy-relevant size.

5. Design Effects

To draw appropriate inferences about the expected effects of a national prospective payment program, our estimated standard errors must reflect the fact that our observations are clustered in a small number of agencies. The variances of the impact estimates generated from standard statistical packages account for the number of patients included in the sample, treating them as though they

are a random sample from an infinite population of patients.[16] However, they do not account for the fact that our sample consists of patients from a limited number of agencies.

To correct for possible nonindependence of observations within agencies, we use SUDAAN software to obtain the appropriate standard errors for our impact estimates. The SUDAAN calculations also account for the greater variance introduced by using sample weights in the regression models.

6. Weighting

As noted previously, we weight observations in the main regression analysis to give each agency equal representation in the analysis. We use this approach for two reasons. First, because the demonstration is implemented at the agency level (not the patient level), the agency is the behavioral unit of interest. Second, the use of weighted data ensures that the impact estimates are not dominated by the experiences of a few large agencies. Preventing this dominance is important because, as shown in Table II.4, the largest agencies have more than 20 times the number of observations as the smallest agencies.

For each patient in agency *i*, we construct the "agency equal" weight (w_i) as the ratio of the average number of patient-level observations per agency (n/k) to the number in the *i*th agency (n_i):

$$(6) \quad w_i = \frac{n/k}{n_i},$$

---

[16] Our sample actually includes the population of (initial) patient admissions taking place in demonstration agencies over the early demonstration period. However, because we wish to make inferences about the outcomes for patients admitted at other times and to other agencies, we treat patients in the data as though they were drawn in a simple random sample from the pool of all (future) patients in all agencies.

where n is the total number of observations for all agencies, and k is the number of agencies. Thus, patient observations for larger-than-average agencies are downweighted, and those for small agencies are given greater weight. For the comparison of impacts between demonstration years, we use a separate agency equal weight in each year. Thus, if an agency is larger (relative to other agencies) in the second year of the demonstration, the weight for observations (patients) in this agency would have a smaller value in the second year than in the first.

7. Robustness Checks

Although we expect the regression models used in the main analysis to be robust, they may be sensitive to four important factors. First, the use of sample weights that equate agencies' representation in the data might give small agencies undue influence in the analysis. Second, the general existence of outliers, either within or across agencies, might have undue influence on the analysis. Third, our main regression models are not strictly appropriate for censored outcome measures and might yield biased estimates of program effects on these outcomes. Finally, other, unobserved characteristics of agencies or their patients might introduce bias into our impact estimates.

a. No Sample Weights

To investigate the sensitivity of our results to the weighting approach, we reestimate the impacts of prospective payment on key outcomes without sample weights. To the extent that the impact estimates are similar under the two approaches, it strongly suggests that our results are not overly influenced by a small number of anomalous observations from agencies with few admissions. Thus, the results may be more broadly interpreted for policy purposes. Conversely, although dissimilar

results under the two sample weights do not necessarily indicate that the main results are "incorrect," they do suggest the need for additional assessment to determine the most valid inferences for policy.

b. Outliers

The distribution of our outcome variables is generally not highly skewed, but it is still important to examine the degree to which our findings on service use are affected by a small number of outlier observations. We use two methods to investigate the effects of outliers. First, to account for outliers at the agency level, we reestimate models excluding all admissions from eight agencies--with the two highest and two lowest mean values for a given outcome in both the control group and the treatment group. Because we use the sample weights that give each agency equal weight, the removal of these eight outlier agencies leaves about 90 percent (71/79) of the weighted sample intact.

In the second approach, we redefine the dependent variable as the natural logarithm of the outcome measure and regress the log variable on the standard set of explanatory variables shown in equation (1). Although this model effectively reduces the influence given to outliers in the simple linear model, it has two major limitations. First, it might substantially understate the true demonstration effects, if the effects are concentrated among those who would have received the greatest number of visits in the absence of the demonstration. Second, the natural logarithm is undefined for values of zero. As a result, we use this model only to examine the effect of outlier observations on the results for total visits and service length.

The log-linear model is given by:

$$(7) \quad lnY = \alpha + X\beta + \delta T + \epsilon,$$

where the variable lnY is the natural log of the outcome variable Y, and the remaining variables and parameters are defined as in equation (1).

To recover the impact of prospective payment on the outcome Y (rather than on lnY) from this model, we must perform the following transformation of the coefficient δ:

$$(8) \quad exp^{\delta} - 1.$$

This transformation yields the impact of prospective payment on the *percentage change* in the outcome Y. To recover its impact on the absolute change in Y (equivalent to our impact estimates from prior models), we multiply this value by the sample mean for the outcome Y.

c. Censoring of the Dependent Variable

The Tobit model is used to obtain consistent estimates of impacts on outcomes where the dependent variable has been censored at a particular value.[17] By comparing the results from these models with those obtained from OLS regressions, we determine the potential sensitivity of our main findings to the effects of censoring.

Variables may be censored from the left, right, or both sides. All variables for the number of visits of a particular type are censored on the left at a value of zero. Variables reflecting the duration of service provision are censored on the right at 366 days, the upper bound of our data.

The Tobit model for our left-censored dependent variables (visits) assumes that an unobserved, underlying index of the need for services (for example, physical therapy visits) exists. If this index (Y^*) exceeds some threshold (zero in our specification), then the patient receives the amount of

---

[17] The consistency of the Tobit model rests on very strong statistical assumptions, and violation of these assumptions may lead to substantial bias in the estimated impacts. Therefore, we prefer to use this model only as a sensitivity test, and to rely on more robust OLS models for the main findings.

services necessary; if the index is less than the threshold, the patient receives no care. The model therefore consists of a probabilistic component for the likelihood of requiring a particular type of service, and a linear component for the expected number of visits, conditional on having a need for services above the threshold:

$$
\begin{aligned}
(9)\qquad & Y^* = \alpha + X\beta + \delta T + \epsilon \\
& Y = Y^* \text{ if } Y^* > 0 \\
& Y = 0 \text{ if } Y^* \leq 0,
\end{aligned}
$$

where Y^* is the underlying index variable that determine the value of the outcome Y, and the remaining variables and parameters are defined as in equation (1).

After the parameters of the Tobit model have been estimated, they are used to obtain predictions of the expected value of the given outcome for each observation. This expected value is given by the product of the probability that a patient uses a given service and the amount of the service used (conditional on it being greater than zero). These predicted values are calculated twice: once assigning a given observation to the control group, and once assigning it to the treatment group. The mean difference between these two predicted values is our estimate of the impact of the demonstration on a given outcome.

The Tobit model to correct for right censoring in the duration of service provision follows the same approach. For our analysis of the patient year, observed duration ranges from 1 to 365 days for complete spells, plus a portion of the sample still receiving services on the 366th day. The model is therefore given by:

$$(10) \qquad \begin{aligned} Y^* &= \alpha + X\beta + \delta T + \epsilon \\ Y &= Y^* \text{ if } Y^* < 366 \\ Y &= 366 \text{ if } Y^* \geq 366. \end{aligned}$$

After the parameters of the Tobit model have been estimated, they are used to obtain predictions as before. The predicted service length (within the first 366 days) is given by the sum of two items: (1) the product of the predicted probability that services end in one year or less and the predicted service duration (conditional on it being less than 366 days); and (2) the product of the probability that services are received for longer than one year and 366, the maximum value observed. These predicted values are estimated twice for each observation, once with the case assigned to the treatment group, and once with it assigned to the control group. The mean difference between these two predicted values is our estimate of the impact of the demonstration on the duration of services within the year after admission to home health care.

d. Unobserved Differences Between Agencies

Given the large treatment-control differences found in selected characteristics (shown in Table II.7), a natural question is whether other, unobserved differences between treatment and control agencies could lead to bias in our impact estimates, because our regressions do not control for these differences. To address this question, we conduct two analyses. In the first, we include the additional set of patient-level control variables available for a subset of patients from the quality assurance (QA) admissions data and reestimate the impacts of the demonstration for key outcomes.[18] In the second, we use a basic difference-in-difference model to estimate selected demonstration impacts; this model is applied to agency-level observations to “net out” any unobserved difference

---

[18] The QA variables, summarized in Table II.3, were not used in the main regression models because they were available for only a subset of patients.

between agencies that remain fixed under the demonstration. Specifically, for a given outcome, the difference-in-difference model calculates the demonstration impact as the average treatment-control difference in the change from the predemonstration period to the demonstration period in agency means for our outcome measures.

Note that if important changes took place in agency characteristics between the predemonstration and demonstration, and these changes were correlated both with agencies' treatment status and the outcome of interest, our estimated impacts based on difference-in-difference models may be biased. While we cannot rule out that these types of changes might have taken place, it seems unlikely, given our available data. For example, perhaps the strongest potential for bias was from treatment agencies altering their case mix in favor of relatively health patients. While such behavior might lead us to overstate the effect of the demonstration through difference-in-difference models, previous findings from Trenholm (1999) provide no evidence of such changes in case mix taking place. These models thus provide a good method for examining the robustness of selected findings.

In practice, the difference-in-difference model amounts to a simple regression model in which the dependent variable is the *agency-level* change in the agency mean for a given outcome (such as total visits per patient) between the predemonstration and demonstration, and the regressors are only a constant term and the agency's treatment status. The coefficient on the treatment-status variable measures the demonstration impact on the outcome of interest.

IV. IMPACT OF PROSPECTIVE PAYMENT ON THE USE OF SERVICES

This chapter presents findings on the impacts of per-episode prospective payment on the home health use of patients admitted to demonstration agencies during the first demonstration year. It begins in Section A with a summary of the impacts we might expect to find, given the incentives of the demonstration payment system. This is followed in section B through D with our main impact findings. As explained in Chapter III, these impact findings are based on patients' service use over the 12-month period after their initial admission to a demonstration home health agency (referred to as a "patient-year"). Section E then presents a largely qualitative analysis of service use within the patient episode. Finally, in Section F, we present findings on the sensitivity and general robustness of our impact findings.

Unless otherwise noted, all counts of patients' home health use reported in this chapter, as well as in Chapters V and VI, have been calculated using statistical weights that give each agency equal representation in our sample. Given this weighting approach, and the potential non-representativeness of agencies participating in the demonstration, the counts of service use reported for control group agencies should *not* be interpreted more generally as national or industry-wide estimates of service use under cost-based reimbursement. (For additional details on our weighting approach and overall analytic approach, see Chapter III.)

A. EXPECTED EFFECTS

A prospectively paid (treatment) agency faces a mixture of incentives in providing services to its patients. Within the first 120 days after a patient admission, a treatment agency is subject to per-episode payment and bears financial risk for all the care that it provides. The agency therefore has a substantial incentive to reduce the number of visits during this "at-risk" period in order to generate

profit. After the at-risk period, however, the incentives to cut visits are less clear. If the agency continues to provide care within the same patient episode, it receives a (prospective) "outlier" payment for each visit. Assuming this payment is less than its per-visit costs, the agency has an incentive to reduce the number of visits to limit losses, although this incentive may be offset by the need to maintain sufficient volume.[1] If the payment is at or above its costs, however, the agency has an incentive to increase the number of visits, rather than lower them. Moreover, after the patient has been discharged, and a 45-day (post-at-risk) gap in care has passed, the agency has the incentive to readmit the patient to receive a second episodic payment.

Despite this mixture of incentives, we expect treatment agencies to reduce their provision of services relative to what they would have been under a cost-reimbursed system. A principal reason is that the incentive to generate savings during the at-risk period should lead to significant reductions in visits during the first 120 days after a patient admission. Moreover, because efforts to reduce visits during the at-risk period may be closely associated with earlier discharge, patients of treatment agencies are likely to receive fewer outlier visits. Thus, even if agencies sought to increase outlier visits for those still receiving care, the net change in the number of these visits per patient may still be negative. In addition, any agency that adopts a policy to reduce visits during the at-risk period probably has limited ability to reverse this policy in the outlier period, given the level of micro-management required. This limitation, combined with the difficulty in determining whether outlier payments are actually above or below costs, suggests that many agencies may lower their outlier visits irrespective of the financial benefits of doing so. Finally, despite an incentive exists to readmit

---

[1] Maintaining volume may be particularly important if, for example, an agency was in danger of having its per-visit costs rise above the limits, or if it had limited flexibility in reducing hours for visiting staff.

patients to obtain a second episodic payment, agencies probably have only limited ability (and desire) to pursue this strategy.

Treatment agencies reported implementing a number of different initiatives to reduce visits.[2] The goal of many of these initiatives was to enable the agencies to reduce the number of visits by discharging patients sooner. Examples cited by several agencies included increasing the use of specialists, for more rapid wound healing; increasing the number of visits early in the episode, to speed healing and learning about self-care routines; placing greater reliance on patients' relatives to assume responsibilities for care; and making more extensive use of hospice care or nursing home care for seriously ill patients. Agencies also identified initiatives to reduce the frequency of visits for a given patient, including improving the supervision of care planning and practices; increasing the use of care maps; substituting services from charitable or other community-based organizations for agency services; and using the telephone, rather than visits, to monitor patients.

Aside from reducing visits for its typical mix of patients, treatment agencies could also cut back on services by developing a patient mix that requires less care than a typical mix. However, in our interviews we found little evidence that treatment agencies replaced their referral sources or altered intake procedures in order to improve their overall patient mix. In fact, most agencies indicated that they had little control over their patient mix, because they could not choose from among existing referral sources and could not easily obtain any new sources, much less ones with a healthier mix. Moreover, the case-mix adjuster mitigates the incentive to attract a healthier mix of patients, because it may result in a lower adjusted payment. (The case-mix adjuster is discussed in Appendix A.)

---

[2] Comments on agencies' behavior under the demonstration are based on site visit interviews with individual agencies conducted near the start of the demonstration, and more recently during agencies' third demonstration year. See Phillips and Thompson (1997) for a summary of the earlier site visit interviews; a report from the more recent interviews is forthcoming.

B. IMPACTS ON SERVICE USE IN THE YEAR AFTER ADMISSION

Prospective payment sharply reduced the average visits per patient during the first 12 months after admission (the "patient-year"). In each month of the patient-year, the average visits provided by treatment agencies were significantly below the average visits by control agencies (Table IV.1).[3] These monthly differences declined steadily but very slowly during the year, ranging from about two and one-half visits in the first two months to just over one visit in the last two months.

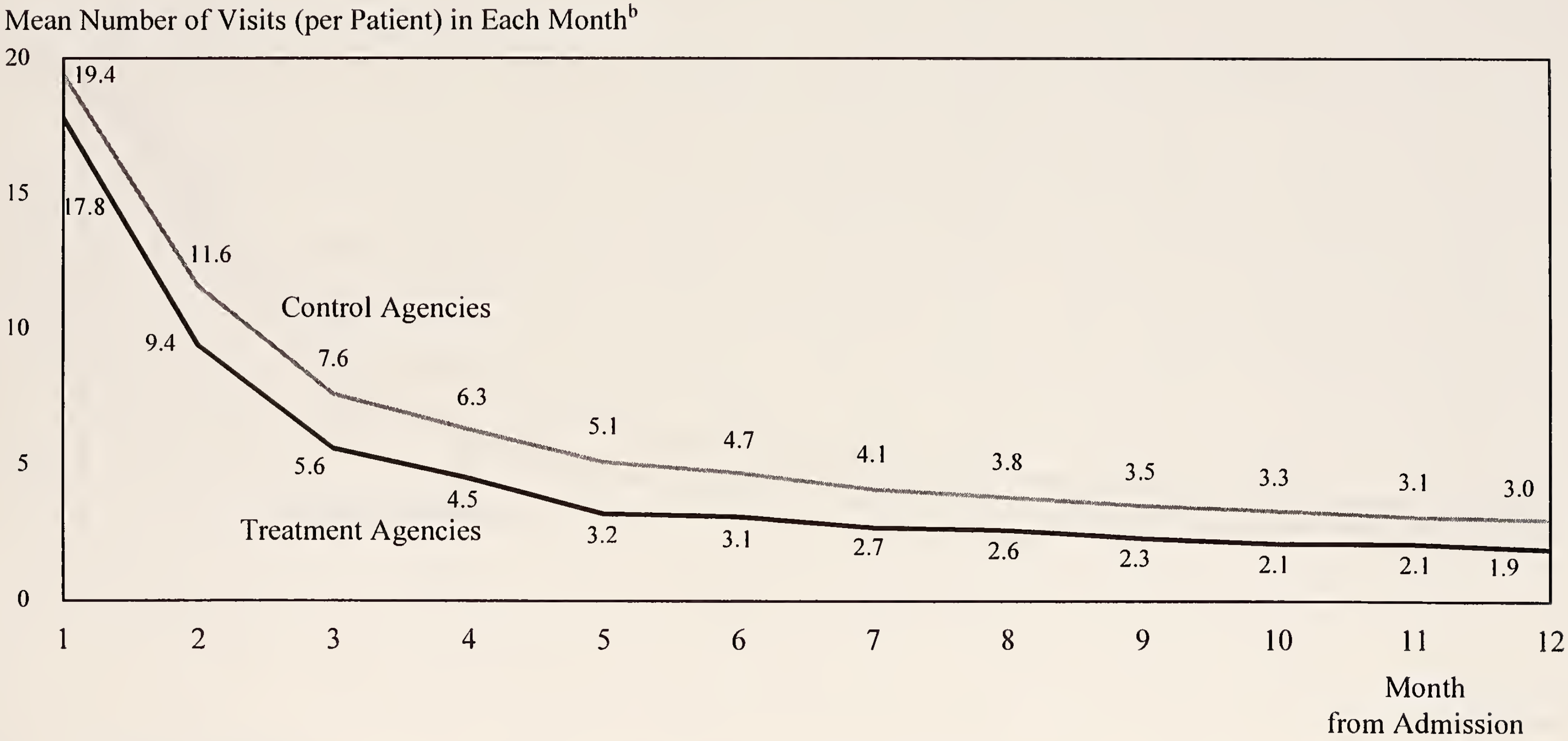
Prospective payment not only lowered the average number of visits received each month but also reduced the likelihood that a patient received any visits after the first month (Figure IV.2). During the second month after admission, for example, 65 percent of patients in treatment agencies received a home health visit, compared with 72 percent of patients in control agencies--a significant difference (impact) of seven percentage points, or 10 percent relative to the rate for the control group. This difference increased to 10 percentage points by the third month and remained at this level through the fifth (first post-at-risk) month, then gradually declined. By month 12, 17 percent of control patients received a visit, compared with only 12 percent of treatment patients--a difference of five percentage points, or roughly 30 percent.

---

[3] The monthly means shown in Figure IV.1 include all patients in our sample, not just those who received care in a given month. In both Figure IV.1 and IV.2 (discussed below), a separate regression was estimated for each month. These estimates include service use provided only by the admitting demonstration agency. When service use from other agencies is accounted for, the gap between treatment and control group agencies closes slightly in most months but remains consistently large and significant.

FIGURE IV.1

AVERAGE MONTHLY VISITS DURING THE YEAR AFTER ADMISSION,[a] BY TREATMENT STATUS



SOURCE: Medicare claims data.

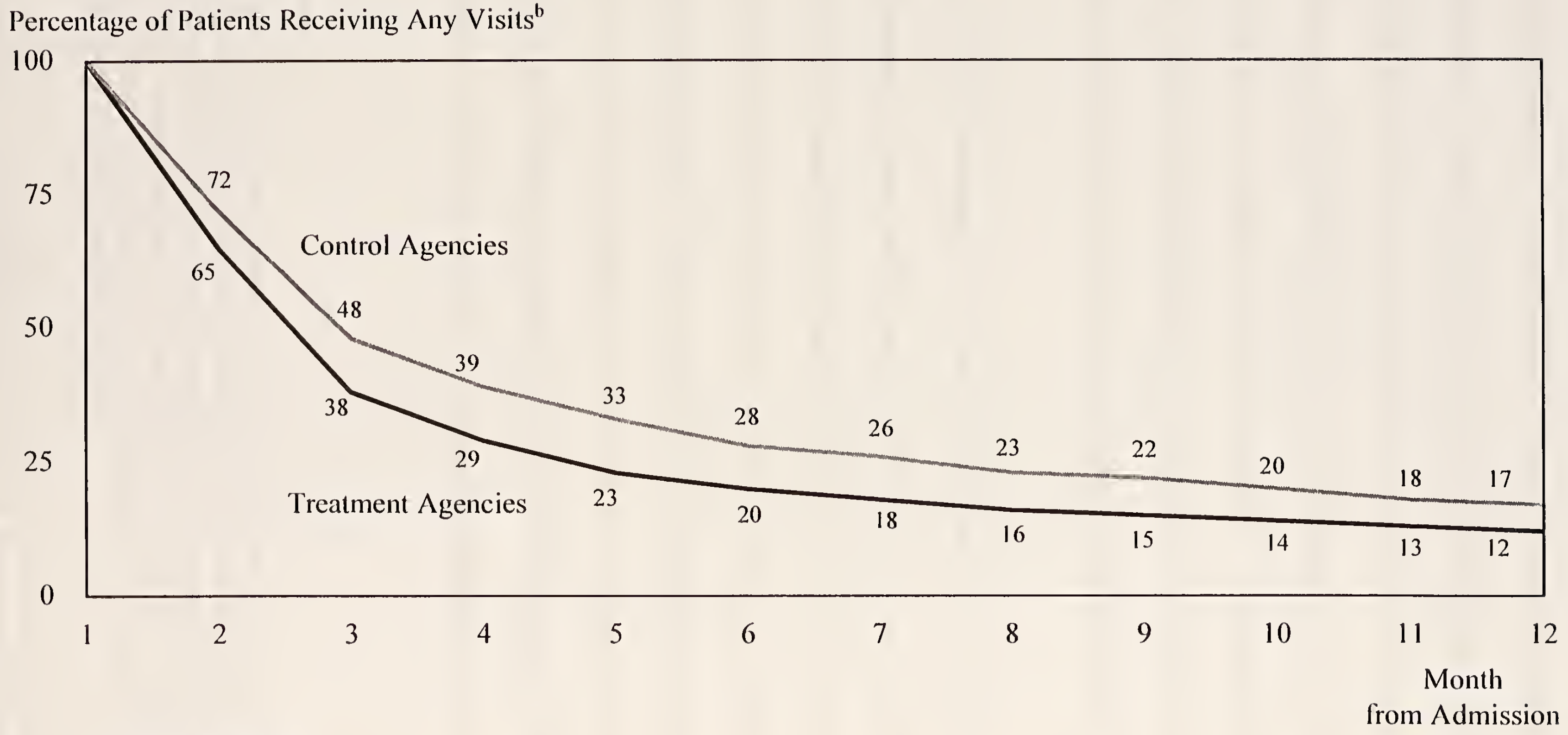
NOTE: Individual observations (patients) weighted to give agencies equal representation in the calculations. Sample size = 26,384 patients in 39 treatment agencies; 25,749 patients in 40 control agencies.

[a] Reflects a patient's first admission to a home health agency participating in the demonstration.

[b] Means for treatment agencies have been regression adjusted (through ordinary least squares models) to control for preexisting differences between patients or agencies. A separate regression was estimated for each month. Each monthly treatment-control difference is statistically significant from zero (p-value < .01), based on a two-tailed test and standard errors inflated to account for the effects of clustering and weighting.

FIGURE IV.2

RATE OF SERVICE RECEIPT IN EACH MONTH FROM ADMISSION,[a] BY TREATMENT STATUS



SOURCE: Medicare claims data.

NOTE: Individual observations (patients) weighted to give agencies equal representation in the calculations. Sample size = 26,384 patients in 39 treatment agencies; 25,749 patients in 40 control agencies.

[a] Reflects a patient's first admission to a home health agency participating in the demonstration.

[b] Means for treatment agencies have been regression adjusted (through logit models) to control for preexisting differences between patients or agencies. A separate regression was estimated for each month. Each monthly treatment-control difference is statistically significant from zero (p-value < .01), based on a two-tailed test and standard errors inflated to account for the effects of clustering and weighting.

[illegible]

[illegible]

[illegible]

As the monthly differences in visits accumulate (Table IV.1), they generate an increasingly large overall program impact.[4] Treatment agencies reduced the average number of visits per patient over the year from about 76 to 58--a statistically significant decline of 18 visits, or 22 percent. More than half the reduction occurred after the first 120 days, when agencies were no longer subject to per-episode payment (unless the patient had a second episode). The magnitude of this post-at-risk impact, coupled with the reduced likelihood of care during a given month, strongly suggests that treatment agencies were able to discharge more patients in the first 120 days than were control agencies.[5]

Although the sizable reduction in visits during the patient-year was largely the result of fewer visits by skilled nurses and home health aides, the use of other services also declined significantly, and the overall visit mix changed little (Figure IV.3). Agencies under prospective payment cut the average number of skilled nursing visits from about 35 to about 26 per patient-year, a reduction of about 9 visits (26 percent) relative to the control group. The reduction in home health aide visits (7.3 visits) was somewhat smaller in magnitude but nearly the same in percentage terms. The decrease in therapy visits was the smallest among the four visit types, as the average number fell from 9.2 in control agencies to 7.8 in treatment agencies, a marginally significant impact of 1.4 visits, or 15

---

[4] As noted previously, these impacts account for services received only from the admitting demonstration agency. When service use from other agencies is accounted for, these impacts are slightly smaller, though still large and highly significant. For example, the impact on total visits in a patient year is -16.3 when including service use from other agencies. This impact is not only still very large and statistically significant, but it is not statistically different from the impact of -18.3 visits based on admitting agency service use alone.

[5] Because a patient who has not received any visits during a month has not necessarily been discharged, a decrease in the probability of a visit in a given month does not prove that the likelihood of discharge has increased. However, such a decline is highly suggestive that earlier discharge has taken place.

TABLE IV.1

CUMULATIVE NUMBER OF VISITS PROVIDED DURING THE YEAR AFTER ADMISSION[a]

Total Visits from Admission Through:	Treatment Group Mean[b]	Control Group Mean	Difference[b] (*p*-value)[c]
At-Risk Period			
30 Days	17.8	19.4	-1.6*** (0.01)
60 Days	27.1	30.9	-3.8*** (0.00)
90 Days	32.8	38.5	-5.7*** (0.00)
120 Days	37.3	44.8	-7.5*** (0.00)
Post-at-Risk Period			
150 Days	40.6	50.0	-9.4*** (0.00)
180 Days	43.6	54.6	-11.0*** (0.00)
240 Days	48.9	62.6	-13.7*** (0.00)
300 Days	53.3	69.4	-16.1*** (0.00)
1 Year	57.6	75.9	-18.3*** (0.00)
Sample Size			
Patients	**26,384**	**25,749**	--
Agencies	**39**	**40**	--

SOURCE: Medicare claims data.

NOTE: Individual observations (patient records) have been weighted to give agencies equal representation in the calculations.

[a] Reflects a patient's first admission to a home health agency participating in the demonstration.

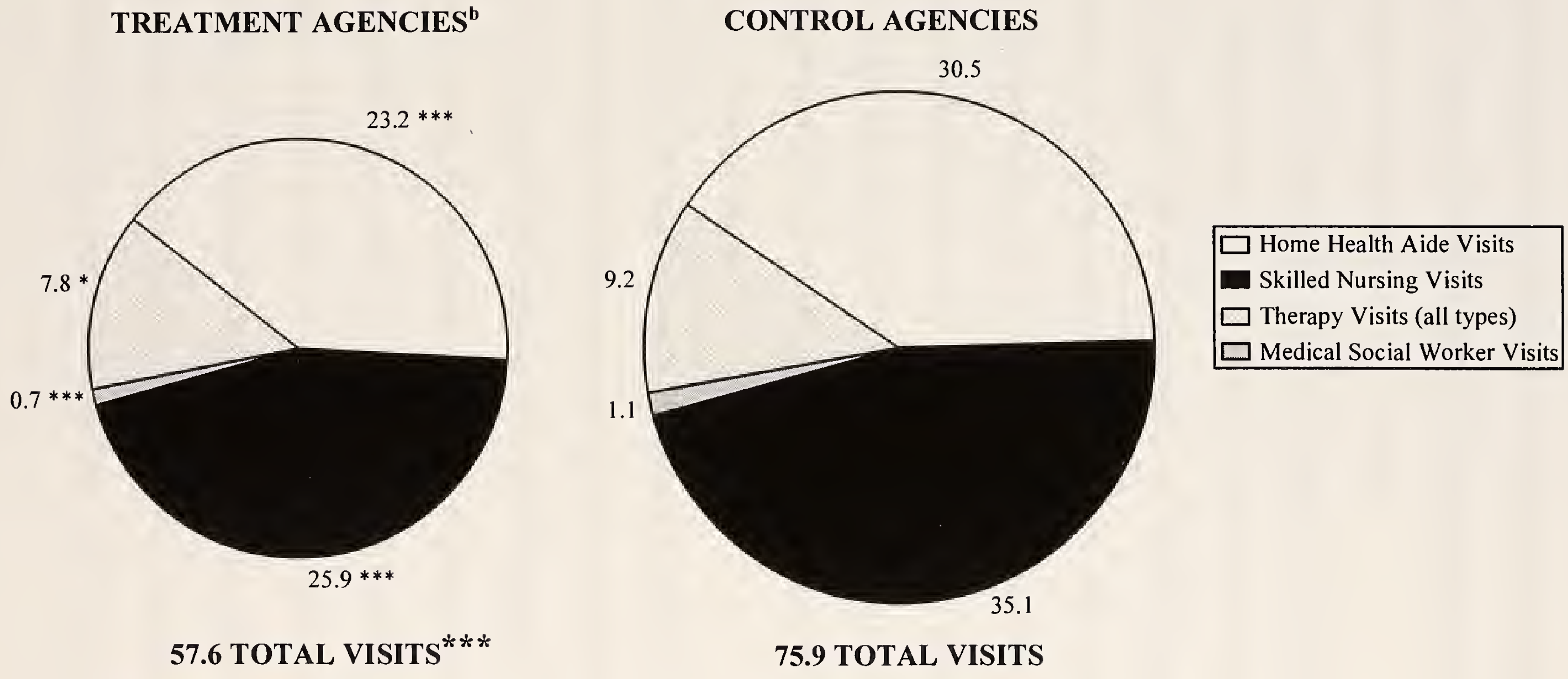
[b] Regression-adjusted (through ordinary least squares models) to control for preexisting differences between treatment and control agencies.

[c] The *p*-value corresponds to a test of whether the treatment-control difference (impact) is statistically different from zero. It is based on standard errors inflated to account for the effects of clustering and weighting.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.
***Significantly different from zero at the .01 level, two-tailed test.

FIGURE IV.3

AVERAGE COMPOSITION OF VISITS IN THE YEAR AFTER ADMISSION,[a] BY TREATMENT STATUS



SOURCE: Medicare claims data.

NOTE: Individual observations (patients) weighted to give agencies equal representation in the calculations. Sample size = 26,384 patients in 39 treatment agencies; 25,749 patients in 40 control agencies.

[a] Reflects a patient's first admission to a home health agency participating in the demonstration.

[b] Regression-adjusted (through ordinary least squares models) to control for preexisting differences between treatment and control agencies. The average number of visits (of each type) for treatment agencies is significantly less than the average number for control agencies (* = p-value < 0.1; ** = p-value < 0.05; *** = p-value < 0.01), based on a two-tailed test and standard errors inflated to account for the effects of clustering and weighting.

percent. In contrast, the impact on medical social workers was highly significant, and the largest in percentage terms. Patients in treatment agencies received an average of 0.7 medical social worker visits during the year, a decline of 0.4 visits, or 36 percent.

The larger decrease in nursing and aide visits compared with therapy visits is consistent with expectations. Many of the initiatives that agencies implemented to cut visits, such as increased reliance on family support and the use of care maps, were targeted largely at services provided by nurses. In addition, if agencies successfully discharged patients earlier, we might expect disproportionate declines in aide visits, as agencies tended to provide a large proportion of these services late in the care episode. Finally, because the availability of therapists is often limited, their services may have been relatively less "oversupplied" by agencies before the demonstration began.[6] Thus, after the demonstration began, treatment agencies may have had less opportunity to make relatively large cuts in therapy visits.

One surprising result is the large decrease in the use of medical social workers. Many treatment agencies indicated that they were planning to make greater use of medical social workers early in the patient episode, in order to speed discharge. We therefore expected to observe medical social worker visits fall only slightly or perhaps even increase. These visits are relatively expensive, however, so the incentive to reduce costs during the at-risk period may have outweighed their perceived benefit in many instances. It is likely that treatment agencies reducing the number of medical social worker visits continued to handle the referrals through nurses or therapists visiting the home, although some treatment agencies may have increased their use of "free" social worker services available through state or charitable organizations. In Illinois, for example, some agencies

---

[6] Some therapists also commented during site visits that because their goals are more measurable than those of other visits staff, they are less likely to overprovide services under cost-reimbursement.

used medical social workers provided by the Illinois Department of Aging. These services would be available to both treatment and control agencies, but only treatment agencies would have a strong incentive to use them to lower costs.

C. IMPACTS ON SPECIFIC SERVICES DURING THE PATIENT-YEAR

The findings discussed in the previous section show how strongly prospective payment affected the total number and general time pattern of patients' visits during the year after the admission. In this section, we investigate these impacts in more detail by examining how the demonstration affected specific services within the patient-year.

1. Impacts During the At-Risk Period (First 120 Days)

Prospective payment reduced average visits during the at-risk period by nearly 17 percent, from roughly 45 visits in control agencies to 37 visits in treatment agencies. As with the impacts for the full patient-year, this decrease most heavily affected nurses, aides, and medical social workers (see Table IV.2). Visits by skilled nurses fell from an average of 21.6 in control agencies to 17.6 in treatment agencies, an impact of 4 visits, or nearly 19 percent. Home health aide visits fell similarly, from 15.5 to 12.8 (roughly 17 percent), and medical social worker visits dropped from 0.8 to 0.5 (nearly 40 percent). By contrast, therapy visits did not decrease significantly during the at-risk period either by type (shown in Table IV.2) or as a group (not shown; *p*-value, 0.38). However, the estimated treatment-control difference in visits by occupational therapists was still quite large (about 20 percent of the control group mean). Thus, it is entirely possible that prospective payment reduced the number of occupational therapy visits during the first 120 days, but we lack the statistical power to detect this effect.

TABLE IV.2

NUMBER OF VISITS PROVIDED DURING THE FIRST 120 DAYS (AT-RISK PERIOD) AFTER ADMISSION,[a] BY SERVICE TYPE

Visit Type	Treatment Group Mean[b]	Control Group Mean	Difference[b] (p-value)[c]
All	37.3	44.8	-7.5*** (0.00)
Skilled Nursing	17.6	21.6	-4.0*** (0.00)
Home Health Aide	12.8	15.5	-2.7*** (0.00)
Physical Therapy	5.2	5.6	-0.4 (0.45)
Occupational Therapy	0.8	1.0	-0.2 (0.20)
Speech Therapy	0.3	0.3	0.00 (0.48)
Medical Social Worker	0.5	0.8	-0.3*** (0.00)
Sample Size			
Patients	**26,384**	**25,749**	--
Agencies	**39**	**40**	--

SOURCE: Medicare claims data.

NOTE: Individual observations (patient records) have been weighted to give agencies equal representation in the calculations.

[a] Reflects a patient's first admission to a home health agency participating in the demonstration.

[b] Regression-adjusted (through ordinary least squares models) to control for preexisting differences between treatment and control agencies.

[c] The *p*-value corresponds to a test of whether the treatment-control difference (impact) is statistically different from zero. It is based on standard errors inflated to account for the effects of clustering and weighting.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.
***Significantly different from zero at the .01 level, two-tailed test.

Although we did not observe a significant drop in mean visits by occupational therapists during the at-risk period, we did find that the likelihood of receiving a visit was reduced significantly (Table IV.3). Specifically, 12.2 percent of patients in control agencies but only 8.3 percent in treatment agencies were seen by occupational therapists during the at-risk period--a difference of nearly four percentage points, or 32 percent. In contrast, the likelihood of receiving a visit by either a physical therapist or speech therapist did not change. Furthermore, no change was observed in the likelihood of receiving any therapy (about 44 percent in both groups; not shown). This finding may suggest that some treatment agencies reduced their "discretionary" use of occupational therapists to supplement the care provided by physical therapists. Indeed, among those receiving therapy services, the likelihood that a patient was seen by both a physical therapist *and* occupational therapist declined significantly, whereas the likelihood of receiving physical therapy alone significantly increased (not shown).

The likelihood of receiving a skilled nursing visit also fell significantly (Table IV.3). This drop is consistent with efforts reported by some treatment agencies to cut (or eliminate) skilled nursing visits to patients referred for therapy services.[7] Despite the decline, treatment agencies continued to provide a skilled nursing visit to more than 93 percent of their patients. Interestingly, we observed essentially no change, significant or otherwise, in the probability of receiving home health aide services. This finding is surprising because home health aide visits are perhaps the most easily substituted for by family members; therefore, we might have expected treatment agencies to eliminate them in some instances. These results suggest that, although treatment agencies were often

---

[7] Before the demonstration, most treatment agencies would have had a nurse perform an assessment visit for all patients referred to the agency to ensure that all skilled care needs would be met. After the demonstration began, however, some treatment agencies had a therapist perform the initial assessment visit, if therapy was expected to be the principal need.

TABLE IV.3

PROBABILITY OF RECEIVING A GIVEN SERVICE DURING THE FIRST 120 DAYS (AT-RISK PERIOD) AFTER ADMISSION[a]

Service	Treatment Group Mean[b] (Percentage)	Control Group Mean (Percentage)	Difference[b] (*p*-value)[c]
Skilled Nursing Visit(s)	93.2	96.6	-3.4*** (0.01)
Home Health Aide Visit(s)	45.2	45.8	-0.6 (0.74)
Physical Therapy Visit(s)	42.4	41.9	0.5 (0.83)
Occupational Therapy Visit(s)	8.3	12.2	-3.9** (0.2)
Speech Therapy Visit(s)	3.0	3.2	-0.2 (.66)
Medical Social Worker Visit(s)	25.5	28.5	-3.0 (0.29)
Sample Size			
Patients	**26,384**	**25,749**	--
Agencies	**39**	**40**	--

SOURCE: Medicare claims data.

NOTE: Individual observations (patient records) have been weighted to give agencies equal representation in the calculations.

[a] Reflects a patient's first admission to a home health agency participating in the demonstration.

[b] Regression-adjusted (through logit models) to control for preexisting differences between treatment and control agencies.

[c] The *p*-value corresponds to a test of whether the treatment-control difference (impact) is statistically different from zero. It is based on standard errors inflated to account for the effects of clustering and weighting.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.
***Significantly different from zero at the .01 level, two-tailed test.

willing to reduce the frequency of aide visits, they were not able to identify suitable resources to eliminate their use altogether.

2. Impacts After the At-Risk Period (120 Days)

As with the impacts in the at-risk period, the overall reduction in visits after the first 120 days was largely concentrated in visits by skilled nurses and home health aides (Table IV.4). In months five through eight, for example, treatment agencies cut an average of 2.9 skilled nursing visits and 2.6 home health visits, which together account for nearly 90 percent of the overall reduction of 6.2 visits during this interval. As percentages of their respective control group means, these impacts on nursing and aide visits are very large--37 and 31 percent, respectively.

One important change from the period comprising the first 120 days is that average therapy visits during the "post-at-risk" period declined significantly. Impacts in months five through eight were particularly large. Relative to control agencies, treatment agencies provided an average of 0.4 fewer physical therapy visits and 0.1 fewer occupational therapy visits--proportional impacts of roughly 33 and 50 percent, respectively. This more evident decline in therapy services after the at-risk period may be the combined result of general initiatives to discharge earlier and more targeted efforts to complete therapy services more quickly. One strategy in particular--concentrating therapy services early in the episode--was mentioned by several agencies and is consistent with an observed decrease in therapy visits only after the at-risk period.

The likelihood of receiving nearly every service dropped significantly after 120 days, strongly suggesting that treatment agencies discharged a larger share of patients during the at-risk period than control agencies (Table IV.5). In months five through eight, roughly 27 percent of treatment agency patients received a visit, compared with 36 percent of control agency patients. This difference

TABLE IV.4

NUMBER OF VISITS PROVIDED AFTER THE FIRST 120 DAYS FROM ADMISSION,[a] BY SERVICE TYPE

	Months 5 Through 8			Months 9 Through 12		
Visit Type	Treatment Group Mean[b]	Control Group Mean	Difference[b] (*p*-value)[c]	Treatment Group Mean[b]	Control Group Mean	Difference[b] (*p*-value)[c]
All	11.5	17.7	-6.2*** (0.00)	8.7	13.3	-4.6*** (0.00)
Skilled Nursing	4.9	7.8	-2.9*** (0.00)	3.3	5.6	-2.3*** (0.00)
Home Health Aide	5.7	8.3	-2.6*** (0.00)	4.7	6.7	-2.0*** (0.01)
Physical Therapy	0.8	1.2	-0.4*** (0.00)	0.5	0.7	-0.2** (0.01)
Occupational Therapy	0.1	0.2	-0.1*** (0.01)	<0.1	0.1	-0.1*** (0.00)
Speech Therapy	0.1	0.1	-0.0 (0.42)	<0.1	<0.1	0.0 (0.64)
Medical Social Worker	0.1	0.2	-0.1** (0.02)	<0.1	0.1	-0.1** (0.01)
Sample Size						
Patients	**26,384**	**25,749**	--	**26,384**	**25,749**	--
Agencies	**39**	**40**	--	**39**	**40**	--

SOURCE: Medicare claims data.

NOTE: Individual observations (patient records) have been weighted to give agencies equal representation in the calculations.

[a] Reflects a patient's first admission to a home health agency participating in the demonstration.

[b] Regression-adjusted (through ordinary least squares models) to control for preexisting differences between treatment and control agencies.

[c] The *p*-value corresponds to a test of whether the treatment-control difference (impact) is statistically different from zero. It is based on standard errors inflated to account for the effects of clustering and weighting.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.
***Significantly different from zero at the .01 level, two-tailed test.

TABLE IV.5

PROBABILITY OF RECEIVING A GIVEN SERVICE AFTER THE FIRST 120 DAYS FROM ADMISSION[a]

	Months 5 Through 8			Months 9 Through 12		
Service	Treatment Group Mean[b] (Percentage)	Control Group Mean (Percentage)	Difference[b] (*p*-value)[c]	Treatment Group Mean[b] (Percentage)	Control Group Mean (Percentage)	Difference[b] (*p*-value)[c]
Any Type of Visit(s)	27.3	36.2	-8.9*** (0.00)	18.1	24.9	-6.8*** (0.00)
Skilled Nursing Visit(s)	25.8	34.8	-9.0*** (0.00)	17.3	24.3	-7.0*** (0.00)
Home Health Aide Visit(s)	14.6	19.5	-4.9*** (0.00)	10.1	13.7	-3.6*** (0.00)
Physical Therapy Visit(s)	7.1	10.1	-3.0*** (0.00)	5.0	6.1	-1.1* (0.06)
Occupational Therapy Visit(s)	1.3	2.6	-1.3*** (0.00)	0.8	1.5	-0.7*** (0.00)
Speech Therapy Visit(s)	0.7	0.9	-0.2 (0.10)	0.3	0.4	-0.1 (0.44)
Medical Social Worker Visit(s)	3.8	7.8	-4.0*** (0.00)	2.8	4.9	-2.1*** (0.00)
Sample Size						
Patients	**26,384**	**25,749**	--	**26,384**	**25,749**	--
Agencies	**39**	**40**	--	**39**	**40**	--

SOURCE: Medicare claims data.

NOTE: Individual observations (patient records) have been weighted to give agencies equal representation in the calculations.

[a] Reflects a patient's first admission to a home health agency participating in the demonstration.

[b] Regression-adjusted (through logit models) to control for preexisting differences between treatment and control agencies.

[c] The *p*-value corresponds to a test of whether the treatment-control difference (impact) is statistically different from zero. It is based on standard errors inflated to account for the effects of clustering and weighting.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.
***Significantly different from zero at the .01 level, two-tailed test.

reflects an impact of about nine percentage points, or 25 percent of the control group rate. In months 9 through 12, the impact was nearly seven percentage points, or about 27 percent.

The likelihood of receiving skilled nursing or home health aide visits experienced the largest percentage point declines in months five through eight; however, the decreases in physical therapy, occupational therapy, and medical social worker visits were even larger as a proportion. For example, the likelihood of a physical therapy visit fell from about 10 percent in the control group to about 7 percent in the treatment group--a proportional impact of 30 percent. This large relative impact is consistent with efforts (noted previously) to concentrate therapy visits early in the episode in order to reach targeted goals more quickly. Interestingly, in months 9 through 12, the impacts on physical therapy weakened substantially and are far smaller than impacts on skilled nurses or aides. Thus, it appears that, for patients with longer-term needs for therapy, treatment agencies had a more difficult time reaching their goals in a shorter period.

3. Summary

Both during and after the (initial) at-risk period, treatment agencies made large cuts in the number of visits by skilled nurses, home health aides, and medical social workers. Because the likelihood of receiving these services also fell sharply after the at-risk period, early discharge appears to be a principal mechanism through which agencies cut back visits.

The reductions in the number of therapy visits during the patient-year were proportionally smaller than for other services, and they took place after the at-risk period. The absence of significant declines in the number of therapy visits during the at-risk period may have been due to increased therapy visits early in the episode, which could have offset cuts later on during the 120-day period. Although the overall probability of receiving therapy services during the at-risk period did not change, the likelihood of receiving a visit from an occupational therapist dropped sharply, suggesting

that treatment agencies sometimes may have dropped occupational therapy as a supplementary service to physical therapy.

D. IMPACTS ON THE DURATION AND TIMING OF SERVICES

Previously discussed results in this chapter have strongly suggested that prospective payment shortens the length of service receipt. As shown in Table IV.6, this inference is in fact true. At any given point after their first admission, a significantly higher percentage of patients in treatment agencies than in control agencies had received their final home health visit from the admitting agency.[8,9] This difference was already about seven percentage points by one month after admission and, by two months, had grown to slightly more than nine percentage points. Thereafter, the differences gradually decrease. By the end of the at-risk period, about 38 percent of patients in control agencies continued receiving home health care, compared with 30 percent of control patients; by one year after admission, these respective rates were about 15 percent and 10 percent.

The pattern of impacts displayed in Table IV.6 strongly suggests that prospective payment did not shorten the length of service receipt for long-term patients only, but instead created a general

---

[8] We define the "last visit" for a given patient as the end date of the last episode *begun within the patient-year*. The "duration (or length) of services" is then given by the number of days from admission until this last visit. For episodes that continue beyond one year, we censor service duration at 366 days.

[9] For two cases, we will understate the actual duration of services by more than the censored value of 366 days; however, neither case should bias the impact estimates in Table IV.6. The first case is if a new episode begins after the patient-year. This case is unlikely to introduce bias because treatment-control differences in new episodes after one year should be trivial. The second case is if we incorrectly assign the end of an episode by day 365 because a 45-day gap was falsely created by missing service data after one year. Such an occurrence should be rare, however, because our four-month data collection "window" (discussed in Chapter II) probably allows 13 or 14 months of nearly complete data for even the latest admissions in the first demonstration year.

TABLE IV.6

TIME FROM ADMISSION UNTIL LAST VISIT [a]

Last Visit Provided by:	Treatment Group Mean[b] (Percentage)	Control Group Mean (Percentage)	Difference[b] (*p*-value)[c]
At-Risk Period			
30 Days	33.0	25.9	7.1*** (0.00)
60 Days	56.6	47.3	9.3*** (0.00)
90 Days	63.4	55.1	8.3*** (0.00)
120 Days	70.4	61.6	8.8*** (0.00)
Post-at-Risk Period			
150 Days	74.0	66.1	7.9*** (0.00)
180 Days	77.4	69.8	7.6*** (0.00)
240 Days	82.5	75.5	7.0*** (0.00)
300 Days	86.0	80.7	5.3*** (0.00)
1 Year	89.6	85.3	4.3*** (0.00)
Average Number of Days from Admission Until Last Visit (Maximum = 366)	110.0	134.7	24.7*** (0.00)
Sample Size			
Patients	**26,384**	**25,749**	--
Agencies	**39**	**40**	--

SOURCE: Medicare claims data.

NOTE: Individual observations (patient records) have been weighted to give agencies equal representation in the calculations.

[a] Defined as the number of days from a patient's first admission to a home health agency participating in the demonstration through the discharge date of the last episode. Due to data constraints, the maximum value is censored at 366 days.

[b] Regression-adjusted to control for preexisting differences between treatment and control agencies.

[c] The *p*-value corresponds (through logit models) to a test of whether the treatment-control difference (impact) is statistically different from zero. It is based on standard errors inflated to account for the effects of clustering and weighting.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.
***Significantly different from zero at the .01 level, two-tailed test.

shift downward for patients of all types. That is, as prospective payment reduced service duration for a given patient from 12 months to 11 months, for example, it reduced it for a second patient from 11 months to 10 months, for a third from 10 months to 9 months, and so on. The result of such a shift, seen clearly in Table IV.6, is a "piling up" of patients who received care for only a short time and a gradual decline in treatment-control differences further from admission.

The treatment-control difference in average service duration, shown at the bottom of Table IV.6, is nearly one month.[10] Specifically, the average time from admission until the last visit was nearly 135 days in control agencies but only 110 days in treatment agencies--a statistically significant difference of 25 days, or 18 percent.

Staff in treatment agencies mentioned two measurable changes to their plans of care near the start of the patient episode: (1) getting a medical social worker in the home as soon as possible; and (2) increasing the number of skilled visits, particularly for therapy patients, in the first few weeks. Ideally, we would use daily or weekly information on patient-service use to investigate the extent of these changes; however, the smallest intervals we could construct from our bill records data is one month.[11] Thus, our analysis of these outcomes focuses on the first month after admission.

Treatment agency patients experienced no significant increase in the likelihood of receiving a medical social worker visit in the first month; in fact, the treatment-control difference is negative (Table IV.7). This result is surprising, as a much larger percentage of treatment agency patients received their last visit in the first month, and it suggests that the incentive to cut medical social

---

[10] Recall that service duration has been censored at 366 days. The actual means are therefore larger, although (as shown in Section F) the estimated program impact is not affected by this censoring.

[11] Beginning in May 1998, all home health agencies were required to identify the date of every service on each bill; previously, however, agencies had to provide the dates of only the first and last service.

TABLE IV.7

TIMING OF SERVICES FROM ADMISSION[a]

Outcome	Treatment Group Mean[b]	Control Group Mean	Difference[b] (*p*-value)[c]
Probability of Receiving a Medical Social Worker Visit in First Month (Percentage)	22.9	24.9	-2.0 (0.46)
Probability of Receiving a Medical Social Worker Visit in First Month, Among Those Who Received a Visit Within a Year (Percentage)	80.0[d]	76.8[d]	3.2 (0.19)
Skilled Nursing Visits in First Month (Number)	9.1	10.2	-1.1*** (0.00)
Physical Therapy Visits in First Month (Number)	2.7	2.7	0.0 (0.86)
Sample Size			
Patients	**26,384**	**25,749**	--
Agencies	**39**	**40**	--

SOURCE: Medicare claims data.

NOTE: Individual observations (patient records) have been weighted to give agencies equal representation in the calculations.

[a] Reflects a patient's first admission to a home health agency participating in the demonstration.

[b] Regression-adjusted (through ordinary least squares and logit models) to control for preexisting differences between treatment and control agencies.

[c] The *p*-value corresponds to a test of whether the treatment-control difference (impact) is statistically different from zero. It is based on standard errors inflated to account for the effects of clustering and weighting.

[d] Sample size: 6,342 treatment agency patients; 8,032 control agency patients.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.
***Significantly different from zero at the .01 level, two-tailed test.

worker visits simply overwhelmed any efforts to use the services of these staff more quickly. This interpretation is further supported when we limit the sample to those who received a medical social worker visit during the patient year. In this case, treatment agency patients were slightly more likely than control agency patients to receive a visit in the first month, but the difference was insignificant and fairly small (about four percent of the control group mean). Thus, while treatment agencies may have provided medical social worker visits more quickly to those that received them, any change of this type was almost certainly small and not as important as the general efforts to reduce or eliminate this service.

Because treatment agencies discharged far more patients during the first month, our inability to count visits on a weekly or daily basis makes it particularly hard to examine whether skilled services were used more intensely near the episode start. Nevertheless, we would not expect to observe a significant reduction in a given service during the first month if agencies were actually providing that service more intensely at the start of the episode. With respect to skilled nursing, this is exactly what we observed. The average number of skilled nursing visits fell significantly during the first month, contradicting the hypothesis that nursing services would be used more intensely. With respect to physical therapy services, however, we found no difference between treatment and control agencies in the number of visits provided during the first month.[12] This null impact certainly does not prove that treatment agencies intensified their use of therapy services near admission. However, given the strong incentive to reduce therapy and other visits throughout the at-risk period, it does suggest that treatment agencies may have done so.

---

[12] The impacts shown in Table IV.7 for both skilled nursing visits and physical therapy visits are based on the full sample and therefore include patients receiving no visits of a given type during the first month. Impacts for the subsample of patients receiving at least one visit are nearly identical (not shown).

E. ANALYSIS OF THE PATIENT EPISODE

To better understand the policy implications of the large declines in service use resulting from prospective payment, we must further analyze service use within the context of the patient episode.[13] Two analyses are particularly relevant. The first is an impacts analysis examining the effect of prospective payment on several outcomes, including the average episode length, the number of episodes per patient, and the number of outlier visits. This analysis provides additional information on how treatment agencies reduced service provision, as well as on whether they "gamed" the payment system (for example, by raising the readmission rate to obtain a second episodic payment). The second is a descriptive analysis of how services were distributed within the initial patient episode under prospective payment, including the distribution of episode lengths, the number of at-risk and outlier visits per episode, and the frequency of visits for patients who continue to receive care. This analysis reveals the type and proportion of services covered by the episodic payment, as well as other periods that might be considered for a nationwide payment system.

Two caveats should be noted with respect to these analyses. First, in order to examine a full year of outlier services, the data cover a 16-month period after admission, rather than the 12-month period used to investigate impacts (discussed in previous sections). Owing to data constraints, this leads to a sample that includes only the first 8 months of demonstration admissions for each agency, rather than the first 12 months. Second, some evidence indicates that treatment agencies' visits per episode decreased during the second year of the demonstration.[14] Our descriptive analysis may

---

[13] Recall that an episode of care, as defined by the demonstration, includes any (at-risk) visits provided during the first 120 days after admission, and any (outlier) visits provided before a 45-day gap in care occurring after the at-risk period has ended.

[14] In Chapter VI, we examine whether the impacts changed between the first and second years of the demonstration. As shown in Table VI.1 of that chapter, in the second demonstration year, the average number of visits during the at-risk period decreased significantly for treatment agency (continued...)

therefore overstate the actual services that prospectively paid agencies provided late in the demonstration.

1. Impacts on Key Episode Characteristics

Given the large impacts on the duration of services, it is no surprise that the average length of the initial episode is much shorter in prospectively paid agencies (Table IV.8).[15] Compared with patients in control agencies, whose initial episode averaged about 131 days, the initial episode for an average treatment patient was more than one month (33 days) shorter.[16] Treatment patients were also only about half as likely as control patients to continue their initial episode more than 16 months after discharge (5 percent and roughly 9 percent, respectively). Although less frequent admission of long-term patients by treatment agencies could have explained this difference, we found no important case-mix differences between the two agency types.[17] Thus, it appears that prospective

---

[14] (...continued)
admissions.

[15] For admissions in the first 8 months, we observed service use for as few as 16 months (480 days). Given this constraint, we censored the lengths of all episodes continuing past 16 months at 481 days.

[16] The impact on average episode length differs from the impact on average service duration (shown in Table IV.6) because the former analysis includes only the initial episode, whereas the latter analysis includes all episodes begun during the patient-year. In addition, the episode-length analysis follows a subsample of patients for 16 months after admission, whereas the analysis of service length follows the full sample for 12 months after admission.

[17] As shown in Table II.5, the observed characteristics of patients in treatment and control agencies were very similar. Moreover, other patient characteristics, obtained from the QA data for selected admissions, were also quite similar between treatment and control agencies (not shown).

TABLE IV.8

CHARACTERISTICS OF THE INITIAL PATIENT EPISODE

Outcome	Treatment Group Mean[a]	Control Group Mean	Difference[a] (*p*-value)[b]
Initial Episode Length (Maximum = 481)	97.8	131.2	-33.4*** (0.00)
Visits in First Month of Outlier Period (5th Month from Admission) for Patients with an Episode Ongoing into the Next Month (Number)	14.3[c]	16.4[c]	-2.1*** (0.0)
Probability of Episode Length > 16 Months (Percentage)	5.0	9.1	-4.1*** (0.00)
Probability of 2nd Episode Within 16 Months (Percentage)	10.7	10.2	0.5 (0.57)
Probability of Receiving Only One or Two Visits (Percentage)	4.3	3.8	0.5 (0.15)
Outlier Visits for Those with at Least One Outlier Visit (Number)	75.8[d]	95.1[d]	-19.3*** (0.00)
Sample Size			
Patients	**20.508**	**22,547**	--
Agencies	**39**	**40**	--

SOURCE: Medicare claims data.

NOTE: Individual observations (patient records) have been weighted to give agencies equal representation in the calculations.

[a] Regression-adjusted (through ordinary least squares and logit models) to control for preexisting differences between treatment and control agencies.

[b] The *p*-value corresponds to a test of whether the treatment-control difference (impact) is statistically different from zero. It is based on standard errors inflated to account for the effects of clustering and weighting.

[c] Sample size: 3,927 treatment agency patients; 5,033 control agency patients

[d] Sample size: 4,926 treatment agency patients; 6,441 control agency patients.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.
***Significantly different from zero at the .01 level, two-tailed test.

payment led treatment agencies to shorten episodes of all different lengths, including those for patients with long-term needs for care.[18]

Treatment agencies provided significantly fewer visits to patients in an ongoing episode, clear evidence that they cut service use through a lower frequency of visits as well as through earlier discharge.[19] For example, during the first month of the outlier period, treatment patients in an ongoing episode received 14.3 visits, compared with about 16.4 visits for control patients (Table IV.8). In fact, we found significant treatment-control differences in ongoing episodes in nearly every other month during the first year after admission, typically ranging between one and two visits (not shown).

We found no significant increase in the rate of second episodes among treatment agencies. Within one year after the at-risk period, treatment agencies readmitted about 10.7 percent of their patients into a second episode, compared with 10.2 percent of patients of control agencies, a statistically insignificant difference. This result is surprising, particularly as the shortening of episodes by treatment agencies "naturally" increased the period over which patients were at risk for readmission. Thus, even if treatment agencies provided care of quality equal to that of control agencies and made no effort to readmit patients to obtain additional episode payments, we still might have expected to observe a significant increase in readmissions. The result therefore suggests that treatment agencies not only did not game the payment system through higher readmissions, but they

---

[18] Additional analyses on the probability of an episode ending in a given month present a pattern similar to that shown in Table IV.6 for service length. This pattern is highly consistent with a shortening of episodes of all different lengths.

[19] We define an initial episode as "ongoing" in a given month if, during the following month, the patient continued to receive services that were part of this episode.

also did not reduce the quality of care to the extent that we would have observed higher readmissions.[20]

In a second type of gaming, consistent with the sharp decrease in average episode lengths, treatment agencies would seek short-stay patients who require only one or two visits during their episodes. Again, however, this increase might not be the result of gaming, because the overall reduction in visits by treatment agencies might lead naturally to an increase in the number of apparent short-stay patients. Results presented in Table IV.8 show that the rate of short-stay episodes was higher in treatment agencies than in control agencies (4.3 percent and 3.8 percent, respectively), but that the difference was modest and not statistically significant. These very low rates of short-stay episodes in both agency groups confirm information from site visits suggesting that referrals of this type were rare. We thus conclude that treatment agencies did not game the payment system by admitting patients with little or no true care needs, perhaps because these patients were difficult to obtain.

The demonstrations offered another possible source of gaming, in which an agency could increase the number of outlier visits to patients in order to raise revenues and, perhaps, to obtain additional per-visit profits. However, for patients who received at least one outlier visit during their initial episode, prospective payment significantly *reduced* the average number of outlier visits for the first 12 months of the outlier period, from about 95 to about 76 (Table IV.8). This outcome is not surprising, as it is difficult to see how treatment agencies could successfully instruct staff to pursue initiatives to reduce visits for only the first 120 days after admission. Moreover, some treatment agencies believed that they were losing money on their outlier visits during the first

---

[20] Of course, poor quality may have led to higher admissions to other home health agencies. We will examine this possibility in a future report on the use of other Medicare-covered services, including home health care by nondemonstration agencies.

demonstration year, and they expressed a desire to reduce the number of these visits. Nevertheless, the size of the reduction (roughly 19 visits) is surprising, particularly since patients in a treatment agency with marginal or limited-care needs might be far less likely to continue their episodes into the outlier period than similar, "low-use" patients in a control agency.

2. Description of the Patient Episode Under Prospective Payment

To better understand how services were distributed under the demonstration version of prospective payment, we describe three key aspects of the patient episode: (1) the distribution of episode lengths, (2) the distribution of services in and between the at-risk and outlier periods, and (3) the rate of service receipt for those with an ongoing episode. Because we observe episodes for a maximum of 480 days (the at-risk period plus a one-year outlier period), our description is incomplete for long-stay episodes that continued beyond day 480. Thus, in an additional analysis, we estimated the average visits per "complete" episode based on assumed rates of completion and service provision for these long-stay episodes.

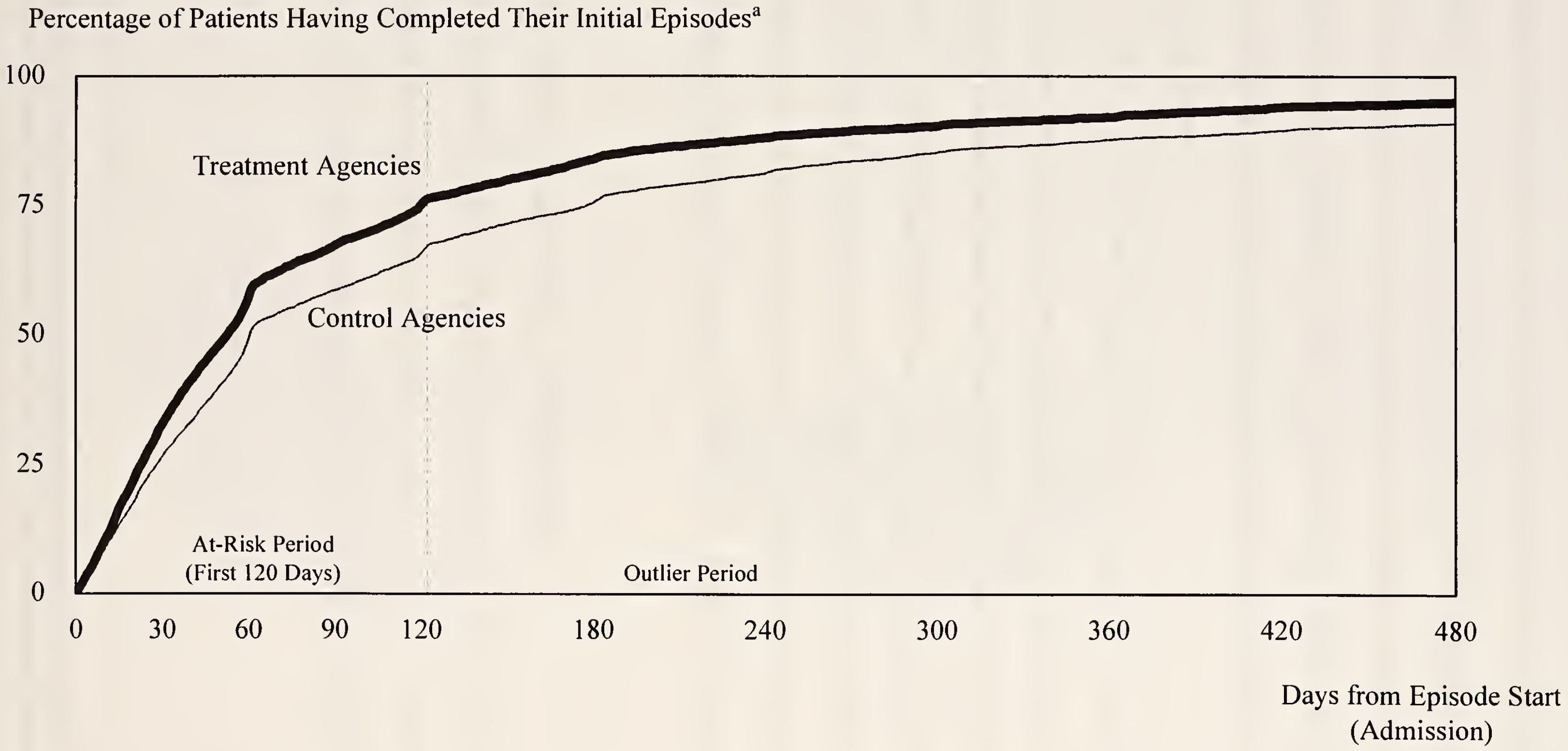
a. Episode Length

We have shown that prospective payment reduced the average length of the censored patient episode by more than one month. In Figure IV.4, we show the distribution of episode lengths for all admissions to treatment agencies during the first eight months of the demonstration. (The distribution for control agencies is also shown, for reference.)

During the first 60 days after admission, episodes in treatment agencies ended at a rate of roughly one percent per day. Although the rate of completion was nearly constant over this period, it increased fairly sharply near the 60th day. This result is somewhat surprising, because only control agencies had to recertify a patient at 60 days; it suggests that treatment agencies followed

FIGURE IV.4

LENGTH OF INITIAL PATIENT EPISODE, BY TREATMENT STATUS



SOURCE: Medicare claims data.

NOTE: Individual observations (patients) weighted to give agencies equal representation in the calculations. Total sample is 22,547 patients in 39 treatment agencies; 20,508 patients in 40 control agencies.

[a] Calculated as weighted averages--no regression adjustment for preexisting differences between patients or agencies.

discharge practices as if the same recertification requirement were still in effect. For the remainder of the at-risk period, the rate at which treatment agencies completed episodes slowed to about half the rate for the first 60 days. Thus, by the end of the at-risk period, roughly three-quarters of all treatment episodes had been completed. Close to the start of the outlier period, a second increase occurred, which probably resulted from the recertification required for all demonstration episodes extending beyond day 120.

During the outlier period, the rate of episode completion gradually and steadily declined. By one year into the outlier period (the maximum length we observed), approximately 95 percent of the episodes in treatment agencies had been completed. We did not observe the completion date for the remaining episodes. However, based on the rate at which episode completion slowed near the end of our observation period, a reasonable estimate is that one-tenth of the remaining episodes would end each 120 days (or about 0.9 discharges per day for patients in this "long-stay" sample). This estimate, which perhaps is upwardly conservative, results in the longest episode lasting roughly 55 months.[21]

b. Distribution of Visits in the At-Risk and Outlier Periods

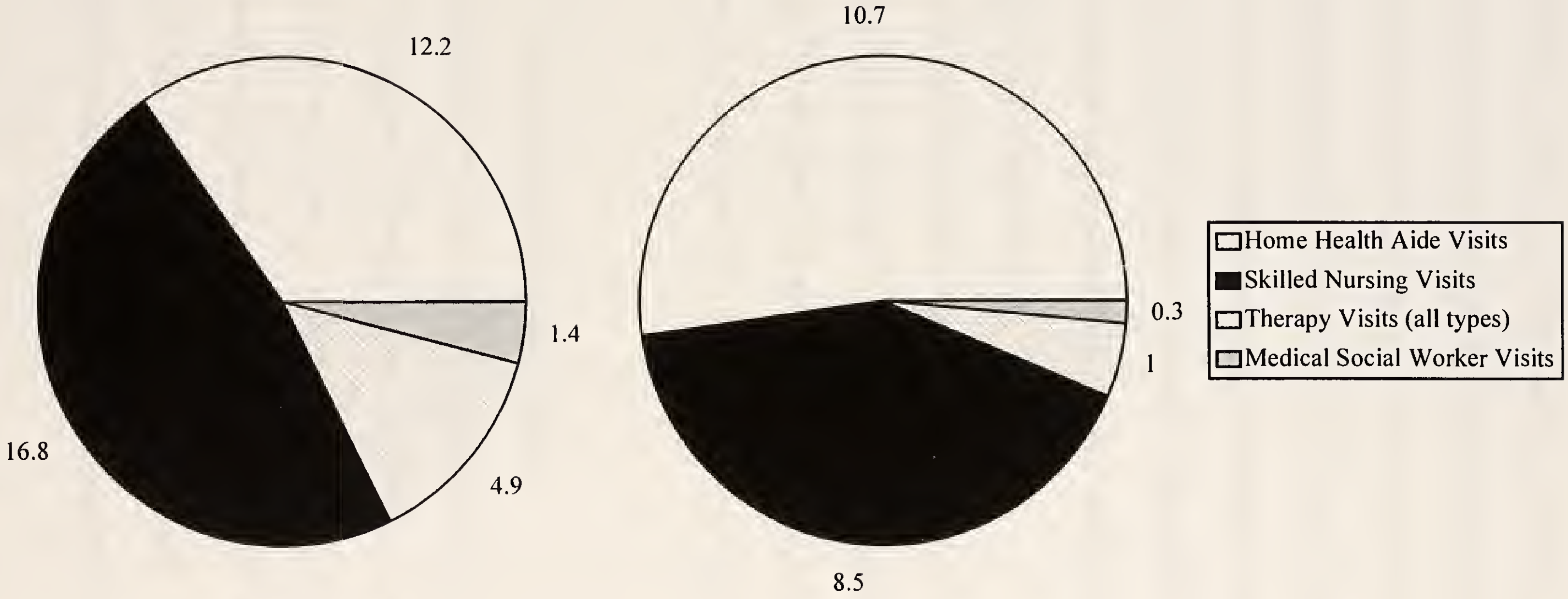
Based on observation of patients' utilization for a full year after the at-risk period, we find that patients in prospectively paid agencies received an average of about 56 visits during their initial episodes (Figure IV.5). This total includes an average of 35.3 visits per patient in the at-risk period

---

[21] Although some episodes that began in the demonstration might have lasted even longer than 55 months, we would expect their number to be very small and to therefore have a trivial effect on the average visits for all demonstration episodes (estimated in Section E.2.d).

FIGURE IV.5

DISTRIBUTION OF SERVICES WITHIN THE INITIAL PATIENT EPISODE, PROSPECTIVELY PAID AGENCIES ONLY



First 120 Days[a]
(At-Risk Period)

35.3 Total Visits per
Patient Episode

One Year After First 120 Days[a]
(Outlier Period)

20.5 Total Visits per
Patient Episode

SOURCE: Medicare claims data.

NOTE: Individual observations (patients) weighted to give agencies equal representation in the calculations. Total sample is 22,547 patients in 39 treatment agencies; 20,508 patients in 40 control agencies.

[a] Calculated as weighted averages--no regression adjustment for preexisting differences between patients or agencies.

[illegible]

[illegible]

and 20.5 visits per patient in the outlier period. Thus, ignoring services provided beyond 480 days, treatment agencies provided about 63 percent of all visits under the fixed per episode payment.[22]

Compared with the outlier period, the distribution of services in the at-risk period was far more heavily weighted toward skilled services. During the first 120 days, nurses provided nearly half (16.8) of all visits, and therapists and medical social workers provided another 17 percent (6.3 visits). The remaining third (12.2) of the visits were provided by home health aides. This distribution differed considerably during the outlier period; home health aides provided more than half of all visits, and nurses provided most of the remainder.

The higher concentration of home health aide services in the outlier period is consistent with the disproportionate provision of aide visits to patients requiring long-term home health care. In total, nearly half the home health aide visits provided per censored episode took place during the outlier period, compared with only one-third of skilled nursing visits and one-fifth of physical therapy visits.

c. Average Visits for Patients with an Ongoing Episode

The preceding summaries of service use during the patient episode have not taken into account discharges taking place during the observation period. In this section, we describe service use for successive months for patients still in care as of the beginning of the month.

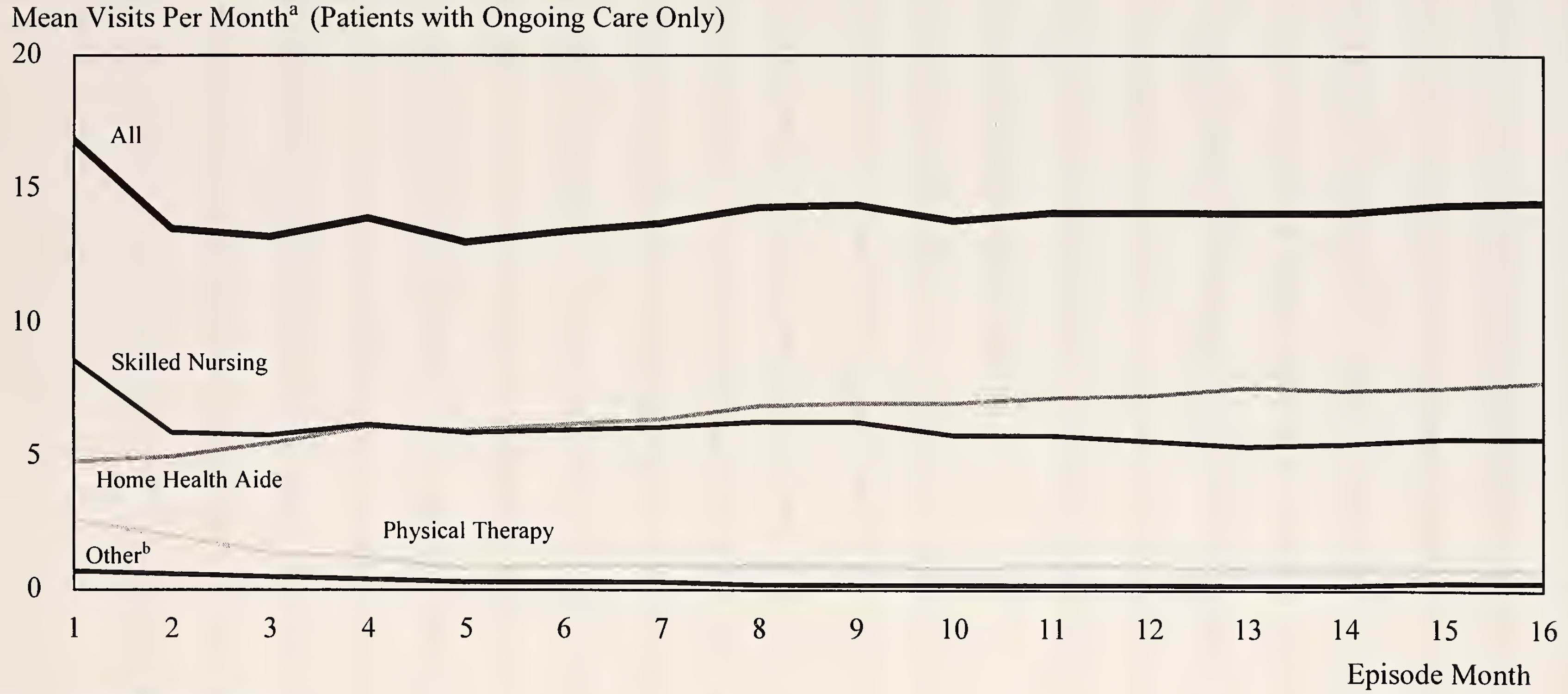
For patients with an ongoing episode, the number of visits provided per month was remarkably stable over time (Figure IV.6). Between the first and second months of the patient episode, the average number of visits per month for an ongoing episode dropped from roughly 17 visits to 14

---

[22] This percentage also ignores "phase-in" patients admitted prior to the demonstration who continued to receive care during the demonstration. Agencies were reimbursed for all phase-in patients on a per-visit basis.

FIGURE IV.6

AVERAGE MONTHLY VISITS AMONG PATIENTS WITH AN ONGOING EPISODE, PROSPECTIVELY PAID AGENCIES ONLY



SOURCE: Medicare claims data.

NOTE: Individual observations (patients) weighted to give agencies equal representation in the calculations. Total sample (at episode start) is 22,547 patients in 39 treatment agencies; 20,508 patients in 40 control agencies.

[a] Calculated as weighted averages for the full sample--no regression adjustment for preexisting differences between patients or agencies. An episode is defined as ongoing in a given month if it has not ended prior to the start of the month.

[b] Includes visits by occupational therapists, speech therapists, and medical social workers.

visits. From months 2 through 16, the number then varied only slightly, ranging from a low of 13 visits to a high of 14.5 visits. This finding suggests that patients requiring longer-term care did not receive a greater intensity of services (albeit of a different mix); rather, they simply received services over a more extended period. However, it also suggests that longer-term patients continued to receive a similar number of visits even many months after admission. Interestingly, the average number of visits for ongoing patients in control agencies was similarly stable over time but also was about one to two visits higher in each month (not shown). This result is consistent with the previously discussed finding that treatment agencies reduced service provision through a lower frequency of visits, as well as through earlier discharge.

The trends by service type supplement the information from Figure IV.5. In the first month of the at-risk period, skilled nurses provided slightly more than half of all visits, home health aides provided about one-quarter, and physical therapists provided most of the remainder. In the second month, the number of skilled nursing visits for ongoing episodes fell sharply, and the other skilled services decreased as well; however, home health aide visits increased slightly. Skilled nursing visits then remained roughly constant for the next several months, while a decline in other skilled services was roughly offset by a rise in the number of home health aide visits.

By the end of the at-risk period, home health aide visits per month had risen to the point that they eclipsed skilled nursing visits for patients with ongoing episodes. The numbers of home health aide and nursing visits subsequently diverged slowly until they leveled off at about seven and five visits, respectively. The mix of visits at that point resembled the distribution for the outlier period shown in Figure IV.5; home health aides provided slightly more than half of all visits, and skilled nurses provided the bulk of the remainder.

d. Estimated Visits per Complete Episode

Because our previous estimates of service use during the outlier period are censored for long-stay episodes that continued beyond 480 days, they understate the actual number of outlier visits provided and overstate the share of visits that fell within the period of the episode payment. Using the data presented in Figures IV.4 through IV.6, however, we are able to reasonably estimate service use patterns for all episodes through their projected date of completion.

Estimating that patients with an episode longer than our 480-day observation period were discharged at a rate of 0.9 per day and received an average of 14 visits per month, the average episode of these "long-stay" episodes included 280 unobserved outlier visits. Because these episodes represent five percent of the total, their additional visits would increase the average for all episodes by about 14 visits. Adding this number to the average of 21 outlier visits for censored episodes (shown in Figure IV.5) leads to a "complete" estimate of 35 outlier visits per episode--by coincidence, the same number as was provided during the at-risk period. Thus, in total, we estimate that prospectively paid agencies provided about 70 visits per episode, split about evenly between the at-risk period and the outlier period.

While this estimate of complete visits per episode is based on key assumptions, it is not likely to understate the number of visits since the assumed episode completion rate is quite low. We therefore view our estimate of 35 outlier visits per episode as a reasonable upper bound, while the estimate of 50 percent of all visits falling in the at-risk period is a reasonable lower bound.

3. Summary

Under the demonstration version of prospective payment, treatment agencies discharged an average patient more than one month earlier from the initial episode of care, compared with control agencies, and also reduced the frequency of services provided in each month. They completed about

three-quarters of all initial episodes within the 120-day (at-risk) period covered by the episodic payment, and completed slightly more than half within 60 days. These rates are much higher than under cost-reimbursement, a further indication that prospectively paid agencies successfully discharged patients earlier.

Agencies displayed no evidence that they gamed the payment system to obtain additional profit. Compared with control agencies, treatment agencies had similar rates of second episodes and of episodes with only one or two visits. In addition, treatment agencies sharply reduced the number of outlier visits for patients receiving any of these visits, contradicting the theory that they would raise the number to obtain per-visit profits.

A small fraction of patients, about five percent, continued their episodes for more than one year after the end of the at-risk period. This rate decreased from about nine percent under cost-reimbursement, strongly suggesting that prospective payment reduced the time until discharge for patients requiring longer-term care. Still, the services these patients required under prospective payment were substantial. Based on reasonable assumptions of episode completion and service use for these longer-term patients, we estimate that about half of all the visits provided in the outlier period were received by these patient more than one year after the at-risk period had ended. Moreover, they accounted for roughly one-quarter of all visits provided to patients during their initial demonstration episode.

The total number of outlier visits per patient is estimated to be about 35. This estimate is nearly equal to the number of visits within the at-risk period, which suggests that the period of the episodic payment covered about half the visits provided per episode.

F. ROBUSTNESS TESTS

In a detailed sensitivity analysis, described in Appendix B, we found no evidence that incomplete billing by some treatment agencies was responsible for the large impacts described in this report. This possibility came to our attention during an early investigation of the billing data, when we identified two treatment agencies with extremely large reductions in their visits per episode during the demonstration. After investigation of the billing data and direct communication with all the agencies, we determined that nine demonstration agencies (eight in the treatment group and one in the control group) had to be dropped from our analysis because of significant evidence of non-billing during the first year of the demonstration.[23] Given lingering concern that other agencies may have had missing bills as well, we conducted a series of sensitivity analyses in which we dropped different combinations of agencies from our sample and reestimated key impacts. The impacts were found to be very similar across these respecifications, strongly suggesting that our results are not the product of missing bill data.

Several additional robustness tests were also performed to ensure that our estimates of demonstration impacts are not heavily dependent on particular observations or statistical procedures. Results from these tests are presented in this section. First, we investigated differences in impacts for selected outcomes (the total number of visits; the number of visits by skilled nurses, home health aides, and physical therapists; and the length of service receipt), using regression models without sample weights. These unweighted estimates place more emphasis on the behavior of larger agencies, rather than place equal emphasis on that of all agencies. Second, we examined whether our impacts changed when we controlled for more extensive differences in the case mix of patients,

---

[23] In Chapter II, we discuss the various reasons that we dropped observations from our analysis sample and note the size of each lost sample.

which are available for a subsample of patients through the data from the QA admissions form (described in Chapter II). Third, we examined the sensitivity of our results to outliers (that is, extreme values for particular patients and agencies). Fourth, we investigated how impacts differed with an analytic model (Tobit) that accounts for censoring of the dependent variable. Finally, we examined whether our results might have been affected by unobserved, agency-level factors not controlled for by our main regression models.

1. Impacts Without Sample Weights

Although giving agencies equal representation in the estimates is the most appropriate approach for our analysis, impacts may be skewed by anomalous mean values of the dependent variable for a few small agencies. The large weight given to admissions in small agencies then makes the impact estimates particularly sensitive to outlier values among these agencies' patients. An alternative approach is to weight observations equally (that is, to use no weights), so that each agency is represented in proportion to its share of patients admitted during the first demonstration year. This approach awards greater influence to agencies with large numbers of admissions than to agencies with few admissions. To the extent that the estimates from the weighted and unweighted data differ, one must be careful when drawing inferences from the previously presented impacts.

The impacts from the unweighted samples are somewhat smaller than those from the weighted data, but they remain very large and quite similar as a proportion (see Table IV.9). For example, the impact on total visits is roughly 15 visits without sample weights, compared with about 18 visits with sample weights. However, because the control group mean for the unweighted data (59.8 visits) is far smaller than the mean for the weighted data (75.9 visits; shown in Table IV.1), the impacts are nearly identical as a proportion of the control group mean (both 24 percent). This pattern holds for other outcomes as well.

TABLE IV.9

IMPACTS OF PROSPECTIVE PAYMENT WITHOUT SAMPLE WEIGHTS

	Unweighted Statistics			Main Analysis
Outcome	Treatment Group Mean[b]	Control Group Mean	Difference[b] (*p*-value)[c]	Impact [b] (*p*-value)[c]
Services by One Year from Admission[a]				
Total Visits	45.1	59.8	-14.7*** (0.00)	-18.3*** (0.00)
Skilled Nursing Visits	20.7	28.4	-7.7*** (0.00)	-9.2*** (0.00)
Home Health Aide Visits	16.5	20.9	-4.4** (0.01)	-7.3*** (0.00)
Physical Therapy Visits	6.1	7.6	-1.5** (0.04)	-1.0 (0.12)
Length of Service (from Admission)[a]				
Days Until Last Visit (Maximum = 366)	102.0	119.6	-17.6*** (0.00)	-24.7*** (0.00)
Probability of Completing Services Within First 120 Days (Percentage)	72.5	66.2	6.3*** (0.00)	8.8*** (0.00)
Sample Size				
Patients	**26,384**	**25,749**	--	**52,133**
Agencies	**39**	**40**	--	**79**

SOURCE: Medicare claims data.

[a] Based on a patient's first admission to a home health agency participating in the demonstration.

[b] Regression-adjusted (through ordinary least squares and logit models) to control for preexisting differences between treatment and control agencies.

[c] The *p*-value corresponds to a test of whether the treatment-control difference (impact) is statistically different from zero. It is based on standard errors inflated to account for the effects of clustering and weighting.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.
***Significantly different from zero at the .01 level, two-tailed test.

The impacts from the unweighted data remain significant and substantial, and they do not suggest that our findings have been distorted by placing disproportionately high weight on selected admissions. Moreover, because the relatively small agencies have relatively high average visits, the proportional impacts for the weighted and unweighted data were very similar. However, the smaller impacts for the unweighted data suggest that relatively small agencies may have made somewhat larger reductions in the number of visits and days of service.[24] In Chapter V, we examine this possibility by estimating impacts for subgroups by agency size.

2. Controlling for Detailed Patient Characteristics

Although we found very few differences in observed patient characteristics between treatment and control agencies (summarized previously in Table II.5), our estimates may be sensitive to unobserved case-mix differences. Fortunately, for a subsample of patients admitted to agencies during the demonstration, detailed data on health status could be obtained from the QA form, which is completed by the home health clinician at admission. From this information, we constructed a number of additional control variables (displayed in Table II.3) that accounted for potential case-mix differences not captured by the patient-level variables used in our main regression models.

The treatment-control differences between most of these variables were small and insignificant, suggesting that our impact estimates have not been biased by unobserved differences between patients (not shown). Regression models that incorporate these variables as additional controls support this inference (Table IV.10). The effects under the QA-enhanced model are smaller than those in our main analysis, but the differences are small. For example, the impacts on total visits

---

[24] Because admissions to smaller agencies receive greater "emphasis" in the weighted sample, the substantial difference in the weighted and unweighted control group means suggests that relatively small agencies provided a higher average number of visits per patient.

TABLE IV.10

IMPACTS OF PROSPECTIVE PAYMENT CONTROLLING FOR MORE DETAILED PATIENT-LEVEL CHARACTERISTICS (FROM QUALITY ASSURANCE DATA)

	Models with Additional Patient-Level Control Variables			Main Analysis
Outcome	Treatment Group Mean[b]	Control Group Mean	Difference[b] (*p*-value)[c]	Impact[b] (*p*-value)[c]
Services by One Year from Admission[a]				
Total Visits	57.6	73.8	-16.2*** (0.00)	-18.3*** (0.00)
Skilled Nursing Visits	26.2	33.7	-7.5*** (0.00)	-9.2*** (0.00)
Home Health Aide Visits	23.4	30.5	-7.1** (0.01)	-7.3*** (0.00)
Physical Therapy Visits	6.1	7.0	-0.9 (0.19)	-1.0 (0.12)
Length of Service (from Admission)[a]				
Days Until Last Visit (Maximum = 366)	111.3	131.5	-20.2*** (0.00)	-24.7*** (0.00)
Probability of Completing Services Within First 120 Days (Percentage)	70.0	63.0	7.0*** (0.00)	8.8*** (0.00)
Sample Size				
Patients	**11,924**	**6,380**	--	**52,133**
Agencies	**39**	**40**	--	**79**

SOURCE: Medicare claims data.

NOTE: Individual observations (patient records) have been weighted to give agencies equal representation in the calculations.

[a] Based on a patient's first admission to a home health agency participating in the demonstration.

[b] Regression-adjusted (through ordinary least squares and logit models) to control for preexisting differences between treatment and control agencies.

[c] The *p*-value corresponds to a test of whether the treatment-control difference (impact) is statistically different from zero. It is based on standard errors inflated to account for the effects of clustering and weighting.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.
***Significantly different from zero at the .01 level, two-tailed test.

differ by only about two visits between the two specifications, and they are very similar as percentages of the control group mean (22 percent in the QA sample, compared with 24 percent in the main sample). Moreover, to the extent that the impacts do differ, it is largely the result of sample differences. Specifically, when the main regression model was rerun on the QA-based subsample, we find almost no differences between the impacts with or without the additional control variables (not shown). Although it is possible that the remaining unobserved characteristics of the treatment and control groups differed, this seems unlikely, given the similarity of the results after controlling for the detailed QA variables. We therefore conclude that unobserved patient characteristics have not biased our main impact estimates.

3. Effects of Outliers

To investigate whether our estimates have been unduly influenced by outliers (at the agency or patient level), we reestimated selected regressions on visits under two alternative specifications. First, we dropped from the regression models the four treatment agencies and four control agencies that had the two highest values and two lowest values of the dependent variable. Second, we respecified the dependent variable as the logarithm.[25]

The estimated impacts of prospective payment under these alternative models are very similar to those found in the main results (see the first two columns of Table IV.11). For each of the outcome measures, the estimated impact changed little in size or statistical significance when outlier agencies were dropped. The effect on total visits is slightly smaller than the main impacts, whereas

---

[25] Because of the shortcomings of the logarithmic specification for dependent variables with many zero values (see Chapter II), we used this specification only to examine impacts on total visits and days of service.

TABLE IV.11

IMPACTS OF PROSPECTIVE PAYMENT ACCOUNTING FOR OUTLIERS AND CENSORING

	Impacts Under Alternative Models			Main Analysis
Outcome	Outliers: Impact with 80 Percent Agency Subsample (*p*-value)[c]	Outliers: Impact with Log-Linear Model (*p*-value)[c]	Censoring: Impact with Tobit Model (*p*-value)[c]	Impact[b] (*p*-value)[c]
Services by One Year from Admission[a]				
Total Visits	-16.6*** (0.00)	-17.4*** (0.00)	[e]	-18.3*** (0.00)
Skilled Nursing Visits	-8.3*** (0.00)	[d]	-7.3*** (0.00)	-9.2*** (0.00)
Home Health Aide Visits	-6.7*** (0.00)	[d]	-4.6*** (0.00)	-7.3*** (0.00)
Physical Therapy Visits	-0.8 (0.15)	[d]	-0.6*** (0.00)	-1.0*** (0.12)
Length of Service (from Admission)[a]				
Days Until Last Visit (Maximum = 366)	-27.5*** (0.00)	-29.8*** (0.00)	-26.6*** (0.00)	-24.7*** (0.00)
Probability of Completing Services Within First 120 Days (Percentage)	9.9*** (0.00)	[d]	[e]	8.8*** (0.00)
Sample Size				
Patients	**Variable**[f]	**52,133**	**52,133**	**52,133**
Agencies	**71**	**79**	**79**	**79**

SOURCE: Medicare claims data.

NOTE: Individual observations (patient records) have been weighted to give agencies equal representation in the calculations.

[a] Based on a patient's first admission to a home health agency participating in the demonstration.

[b] Regression-adjusted (through ordinary least squares and logit models) to control for preexisting differences between treatment and control agencies.

[c] The *p*-value corresponds to a test of whether the treatment-control difference (impact) is statistically different from zero. It is based on standard errors inflated to account for the effects of clustering and weighting.

[d] Outcome is not estimated by log-linear model because of shortcomings of estimating this model with zero values.

[e] Outcome is not censored.

[f] Total sample sizes range from 41,742 to 49,377, depending on the agencies dropped for a given outcome.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.
***Significantly different from zero at the .01 level, two-tailed test.

the impact on days of service is slightly larger. The impacts under the log specifications are also similar, particularly for total visits, which differ by less than one visit.

4. Effects of Censoring

Several of our findings may be subject to bias resulting from censoring of the dependent variable.[26] All the outcomes associated with the number of visits (by type) reflect cases of left-hand censoring (at zero), whereas the length of service reflects a case of right-hand censoring (at day 366). The statistical properties of these two forms of censoring are closely related, and, in principle, Tobit models control effectively for the possible bias introduced by both.[27]

The estimates from the Tobit models suggest that censoring does not explain the significant impacts on the number of visits for selected services and the length of service (see the third column of Table IV.11). Comparing the Tobit and main results shows that the statistical significance and direction of all the impacts are consistent, and that their magnitudes are similar. (The only difference of any size is that for the estimated impact on home health aide visits: -7.3 visits under the OLS regression and -4.6 under a Tobit regression.) Impacts on other service types under Tobit regressions were similar to those discussed previously (not shown).

---

[26] Under censoring, the dependent variable reflects a combination of discrete and continuous variables. Tobit models account for this censoring by estimating the parameters of a single model that predicts both the probability of using a service and the expected amount of the service.

[27] In the main part of the chapter, we report the prospective payment impacts using OLS specifications, because the statistical properties of the Tobit model can make its estimates unreliable. To the extent that the estimated impacts using Tobit and OLS are similar, it strongly suggests that our findings are not affected by the censoring or truncation of the dependent variable.

5. Influence of Unobserved Variables

A final issue is whether our results might have been affected by unobserved differences between treatment and control agencies that we cannot control for through our demonstration-based regression models. For example, if unobserved factors affecting service use between the predemonstration and demonstration (such as local market conditions, quality of management, or patient case mix) differ between treatment and control agencies, then our estimated impacts may be biased.[28]

To investigate the influence on our findings of unobservable factors at the agency level, we estimated demonstration impacts on selected outcomes as the treatment-control difference in the change from the predemonstration period to the demonstration.[29] This approach, commonly referred to as a difference-in-difference model, nets out any fixed differences between agencies that could introduce bias into our impact estimates. Because our data from the predemonstration period is limited to the first 120 days after admission, our sensitivity tests focus only on outcomes within the at-risk period.

For two key outcomes, total visits and days of service receipt (both over the 120-day period from admission), the impacts under the difference-in-difference models are very similar to the main impacts (Table IV.12). With respect to visits, for example, estimates from the difference-in-difference model indicate that agencies under prospective payment cut eight visits per patient during the at-risk period, slightly more than the estimate from the main regression model. Interestingly,

---

[28] See Chapter III for a more complete discussion of this issue. While the random assignment of agencies into the treatment and control groups should substantially reduce the likelihood of such unobserved differences, our samples of agencies are small enough where such unobserved differences may exist.

[29] The predemonstration data reflect all episodes that began during the first eight months of the fiscal year preceding an agency's entrance into the demonstration.

TABLE IV.12

IMPACTS OF PROSPECTIVE PAYMENT BASED ON THE CHANGE FROM PREDEMONSTRATION TO DEMONSTRATION (DIFFERENCE-OF-DIFFERENCE MODEL)

	Difference-of-Difference Model			Main Analysis
	Treatment Group Mean	Control Group Mean	Difference (p-value)[c]	Impact[b] (p-value)[d]
Total Visits During First 120 Days from Admission[a]				
Demonstration	34.6	44.8	-10.2	
Predemonstration	43.4	45.7	-2.3	
Difference	-8.9	-0.9	-8.0*** (0.00)	-7.5*** (0.00)
Days Until Last Service Within First 120 Days from Admission)[a]				
Demonstration	65.4	73.3	-7.9	
Predemonstration	70.6	72.6	-2.0	
Difference	-5.2	0.7	-5.9*** (0.00)	-9.3*** (0.00)
Sample Size				
Patients (Demonstration)	**e**	**e**	--	**52,133**
Patients (Predemonstration)	**e**	**e**	--	--
Agencies	**40**	**39**	--	**79**

SOURCE: Medicare claims data.

NOTE: Agencies have been given equal representation in the calculations.

[a] Based on a patient's first admission to a home health agency participating in the demonstration.

[b] Regression-adjusted (through ordinary least squares models) to control for preexisting differences between treatment and control agencies.

[c] The p-value corresponds to a test of whether the treatment-control difference between the demonstration and pre-demonstration is statistically significant from zero.

[d] The p-value corresponds to a test of whether the treatment-control difference (impact) is statistically different from zero. It is based on standard errors inflated to account for the effects of clustering and weighting.

[e] The patient sample sizes are as follows: 26,384 control group and 25,749 treatment group during the demonstration (same as the main analysis); and 22,260 control group and 21,849 treatment group during the predemonstration. Note that these sample sizes are less relevant to the difference-in-difference models because these models use agency (not patient) averages to arrive at the treatment and control group means in each period. The statistical precision of these models is thus a function of the number of agencies, rather than patients, in the sample.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.
***Significantly different from zero at the .01 level, two-tailed test.

control agencies exhibited almost no change in visits per patient between the two periods, further highlighting the substantial decreases that prospective payment caused treatment agencies to make.

6. Overview of Robustness Checks

In conclusion, the various robustness checks confirm the stability of our impact estimates. The estimated effects of prospective payment are large regardless of the how we weighted the sample or what statistical model we used. The findings are not driven by incomplete data on visits for treatment group agencies, by a few agencies with exceptionally high or low visits during the period, or by unobserved differences between agencies or their patients.

G. SUMMARY

Based on a thorough analysis of admissions in the first year of the demonstration, we find very strong and consistent evidence that prospective payment generated substantial, broad-based reductions in the number and duration of services provided to patients.

The key impacts on service use are as follows:

- Prospective payment reduced the average number of visits provided to a home health patient during the first 12 months after admission, from about 76 visits to 58 visits, a decrease of 18 visits, or 24 percent.

- The number of visits for each category of home health service--skilled nursing, home health aide, therapy (all types combined), and medical social worker--dropped as a result of prospective payment. During the year after admission, these decreases ranged from 19 percent to 32 percent.

- Within the first 120 days after admission (the at-risk period), prospective payment reduced the total visits per patient from about 44.8 to 37.3, a decrease of 7.5 visits, or 17 percent. Every service type but therapy contributed to this decrease.

- From the 5th through 12th month after admission, the average number of visits fell from 31.0 to 20.2, an even greater drop than during the at-risk period. All services contributed to this decrease.

- All impact estimates were highly robust across an extensive set of sensitivity tests.
- Early discharge was a key reason for all decreases in service use.
- Reductions in the frequency of visits also contributed to these decreases.
- There is no evidence that treatment agencies gamed the payment system in order to obtain additional profit.

The descriptive data indicate that:

- Three-quarters of all treatment agency patients completed their initial episodes within 120 days (the at-risk period). More than half completed their episodes within 60 days.
- Approximately half of all visits provided to the average treatment agency patients were covered by the episodic payment.
- Based on reasonable assumptions of service use by patients whose episode extended beyond our observation period, roughly one-half of all outlier visits occurred more than 16 months after admission and were directed to a small fraction of longer-term patients.

These findings provide valuable information to policymakers who are considering the creation of a nationwide program of prospective payment. In additional analyses, presented in Chapter V, we examine differences in impacts by key agency- and patient-level subgroups. In Chapter VI we then present an analysis of whether key demonstration impacts changed during the second year of the demonstration.

V. SUBGROUP IMPACTS

Agencies may have responded differently to the demonstration incentives depending on their individual characteristics, or they may have responded differently for certain patient types. To investigate these possibilities, we reestimated the key overall demonstration impacts presented in Chapter IV for a series of agency- and patient-level subgroups most likely to be associated with differences in agency behavior.

A key issue for the agency subgroup analysis is whether the reductions in service use observed for the entire group of demonstration agencies were evident for specific agency types. A finding that impacts were significant across most or all agency subgroups strengthens the validity of the main findings and suggests that they may be more easily applied to agencies outside the demonstration. A second issue is whether demonstration impacts differed between specific subgroups of agencies (for example, between nonprofit and for-profit agencies). To the extent that differences existed, it suggests that prospective payment offered agencies varying opportunities (or incentives) to achieve savings and perhaps profitability, an important consideration when implementing a nationwide payment system.

In the patient-level subgroup analysis, we are principally interested in identifying which types of patients, if any, experienced the deepest reductions in service use under a system of prospective payment. Having identified such patients, we can then focus on them in later reports that examine (for example) patient outcomes, the use of other Medicare-covered services, and patient retention and selection.

For each agency and patient subgroup (detailed in Section A), we estimated program impacts on several key measures of service use and duration. We then conducted statistical tests to determine

(1) whether the program impact for a given subgroup was significantly different from zero; and (2) whether a significant difference existed in the impacts between subgroup pairs (for example, between for-profit and nonprofit agencies).

A. EXPECTED IMPACTS AMONG SUBGROUPS

In this section, we define the specific subgroups that were investigated and discuss the possible differences that we expected to find between subgroups.

1. Agency Characteristics

The five agency-level characteristics that we examined were the agencies' (1) for-profit status, (2) predemonstration practice patterns, (3) size, (4) auspice, and (5) base-year costs. From these respective characteristics, we defined the following subgroup pairs:

- Proprietary or nonprofit
- "High-use" practice pattern or "low-use" practice pattern. These subgroups are defined from a case-mix-adjusted index of agencies' service provision in the base year. (See Chapter II for details). High-use agencies have an index value above the median value; low-use agencies are at or below the median value.
- Small size or large size. Small agencies provided fewer than 30,000 visits in the base year; large agencies provided more than 30,000 visits.
- Hospital-based or freestanding
- Below the cost-limits or above the cost-limits. These subgroups are defined from a comparison of each agency's base-year cost per visit relative to the Section 223 limits.

Proprietary agencies had a clear incentive to decrease visits under the demonstration, but nonprofit agencies had a combination of incentives. Like for-profit agencies, nonprofits had an incentive to reduce visits and generate a surplus that could be applied toward charity work, capital improvements, or other objectives. However, they might have resisted this incentive because their

mission was focused more exclusively on the well-being of patients or informal caregivers. Moreover, in light of the loss protection under the demonstration, some nonprofit agencies might have used prospective payment as an opportunity to increase services to their patients. On average, therefore, prospective payment might have resulted in considerably smaller reductions in visits by nonprofit agencies than in those by for-profit agencies.

It is likely that agencies exhibiting high-use practice patterns before the demonstration began faced less pressure from managed care organizations (and other market forces) to control their service provision under cost-based reimbursement. Because these agencies made less effort to control prior utilization, they probably had the opportunity to make larger-scale reductions in visits before adversely affecting patient outcomes when compared with low-use agencies. Thus, when faced with the incentives of prospective payment, high-use agencies probably made larger reductions in services than did their low-use counterparts.

Smaller agencies (those with fewer than 30,000 visits per year) might have responded differently to the demonstration incentives than did large agencies, but the direction of this difference is uncertain. On the one hand, smaller agencies might have found it easier to communicate with their staff and implement changes in care patterns quickly, thereby achieving large reductions in service use, at least during the first year of the demonstration. On the other hand, these agencies might have been more hesitant about making large cuts because any losses in volume could have generated sharp increases in per-visit costs and loss of staff.

Hospital-based agencies have a more general incentive to maximize profits for the entire hospital system, so their response to the demonstration might have differed from that of freestanding agencies. If its principal concern was patient or physician satisfaction with the hospital system, then a hospital-based agency might have largely ignored the incentives of the demonstration to reduce

the number of visits. In addition, such an agency might have simply had little interest in responding to demonstration incentives because its revenues represent only a small fraction of the revenues of the entire system. However, a hospital-based agency also has advantages that could have allowed it to respond more aggressively to the demonstration. For example, because it has more direct access to referrals, it might have been less concerned with large losses in volume that could result from substantial cuts in visits per patient. The direction of any differences between hospital-based and freestanding agencies is thus uncertain.

Prospective payment provided agencies above their base-year cost limits a very different incentive than they faced under cost-based reimbursement. In the absence of the demonstration, an agency above the limits had an incentive to increase visits per patient in order to raise its overall volume and gain control over its per-visit costs. Under prospective payment, however, its principal goal was the same as if it were below the limits (that is, to reduce the cost per episode by providing fewer visits or by lowering the cost per visit). Therefore, prospective payment might have led to greater treatment-control differences (impacts) for agencies above the limits because (1) treatment agencies in this subgroup cut more visits, and (2) control agencies in this subgroup raised the number of visits.

2. Patient Characteristics

The three patient characteristics that we examined were the patients' (1) ability to take oral medications, (2) access to informal caregivers, and (3) expected costs at admission. The respective subgroup pairs that we defined from these characteristics were as follows:[1]

---

[1] All the information used to form these patient characteristics subgroups has been obtained from the quality assistance (QA) forms that the agencies completed at admission. Because these data are available for only a subsample of patients, the sample for this analysis includes less than half the (continued...)

- Independent in taking oral medications or not independent in taking oral medications
- Other caregiver available or no other caregiver available. Patients are assumed to have other care available if they receive assistance from paid help or from another person residing in the home, or if they reside in an assisted-living facility.
- High expected costs or low expected costs. A patient has high expected costs if he or she falls in the top quartile of predicted costs for all patients in our sample.[2]

Because patients who have other caregivers in the home might have offered agencies an opportunity to substitute for their own care or to discharge earlier (or both), we expect that treatment agencies made larger cuts in visits and service duration for this subgroup than for patients without available caregivers. Treatment agencies might have made large reductions in visits by home health aides in particular, because they might have been able to substitute assistance from other caregivers for aide visits throughout the episode of care.

Patients who cannot take oral medications independently might have a greater need for education and a more severe condition than those who are able to take these medications independently. For these reasons, treatment agencies might have been more reluctant to make large-scale reductions in these patients' service use. We therefore expect that the decrease in the number of visits was more modest for patients who could not take their oral medications independently.

---

[1] (...continued)
number of admissions used in the main analysis (and the evaluation of agency-level subgroups). In addition, we dropped all the observations for nine agencies from the analysis because their QA-based subsample includes only a very small number of admissions (fewer than 20). We expect that these changes to the sample will have little effect on the accuracy of our results; as we note in the sensitivity analysis discussed in Chapter IV, the main impacts differ only slightly when using this QA-based subsample. However, the precision of our estimates will be weaker than it would have been had we used the full sample

[2] The variables used to calculate predicted costs are those identified from Phillips et al. (1992) as increasing costs for the average patient by more than 30 percent.

Patients with high expected costs had characteristics that place their estimated resource use in the top quartile among all patients in the sample. Because these patients traditionally received a very high number of home health visits, they presented a potentially major financial burden to agencies that were subject to per-episode payment. We therefore expect that treatment agencies tried to reduce the number of visits for these high-cost patients by more than the reduction for other patients. Moreover, to the extent that inefficiencies in the care provided by agencies exist, such inefficiencies might have been largest for patients in the high-cost subgroup, making it easier for agencies to reduce services for this group. Because early discharge may be a key strategy to cutting visits for patients with high expected costs, prospective payment might have led to a particularly large cut in service duration for this subgroup.

B. IMPACTS FOR AGENCY SUBGROUPS

As expected, relative to those with low-use practice patterns, agencies with high-use practice patterns did make larger cuts in visits in response to the demonstration (Table V.1). Although both agency types significantly reduced total visits in the patient-year, high-use agencies cut about twice the number of visits as did low-use agencies. The key factor driving this overall difference was a significantly larger reduction in home health aide visits by high-use agencies. This finding might have been expected as high-use control agencies provided far more aide visits than did the low-use control agencies, suggesting that high-use agencies had a substantial opportunity to reduce aide visits after an incentive became available. It thus supports the general hypothesis that high-use agencies made relatively larger cuts in services simply because they had more "room" to do so.

Surprisingly, nonprofit and for-profit agencies responded very similarly to the demonstration. (Table V.2). Both agency types generally made large and significant reductions in the quantity and duration of services. In fact, the only significant difference between the two subgroups was the

TABLE V.1

IMPACT OF PROSPECTIVE PAYMENT ON SERVICE USE, BY WHETHER THE AGENCY HAS A HIGH-USE OR LOW-USE PRACTICE PATTERN[a]

	Treatment Group Mean[c]	Control Group Mean	Difference[c] (p-value)[d]
Service Use (In the Year After Admission)[b]			
Total Visits			
High-use practice pattern	67.6	91.2	-23.6***
Low-use practice pattern	38.7	50.4	-11.7***
Subroup difference			-11.9** (0.03)
Skilled Nursing Visits			
High-use practice pattern	29.5	41.4	-11.9***
Low-use practice pattern	18.2	24.6	-6.4***
Subroup difference			-5.5 (0.14)
Home Health Aide Visits			
High-use practice pattern	28.8	39.4	-10.6***
Low-use practice pattern	13.0	15.7	- 2.7**
Subroup difference			-7.9** (0.05)
Physical Therapy Visits			
High-use practice pattern	6.8	7.5	-0.7
Low-use practice pattern	6.2	7.5	-1.3**
Subroup difference			0.6 (0.72)
Service Length (Through One Year After Admission)[b]			
Days Until Last Visit (Maximum = 366)			
High-use practice pattern	115.0	146.0	-31.0***
Low-use practice pattern	97.7	115.7	-18.0***
Subroup difference			-13.0* (0.09)
Probability of Last Visit by Day 120			
High-use practice pattern	68.1	57.7	10.4***
Low-use practice pattern	74.5	67.2	7.3***
Subroup difference			3.1 (0.45)

	High-Use Practice Pattern		**Low-Use Practice Pattern**	
Sample Size	**Patients**	**Agencies**	**Patients**	**Agencies**
Treatment Group	**14,218**	**17**	**12,166**	**22**
Control Group	**9,689**	**25**	**16,060**	**15**

SOURCE: Medicare claims data.

NOTE: Individual observations have been weighted to give agencies equal representation in the calculations.

[a] An agency is defined as having a high-use practice pattern if its (case-mix-adjusted) number of visits per episode in the base (predemonstration) period was above the median for all agencies in the sample. Otherwise, it is defined as having a low-use practice pattern.

[b] Reflects a patient's first admission to a home health agency participating in the demonstration.

[c] Regression-adjusted (through ordinary least squares and logit models) to control for preexisting differences between treatment and control agencies.

[d] p-value corresponds to a test of whether the subgroup difference is statistically different from zero. It is based on standard errors inflated to account for the effects of clustering and weighting.

*p-value < .10, two-tailed test, corresponding to a test of whether the treatment-control difference within subgroup is different from zero.
**p-value < .05, two-tailed test, corresponding to a test of whether the treatment-control difference within subgroup is different from zero.
***p-value < .01, two-tailed test, corresponding to a test of whether the treatment-control difference within subgroup is different from zero.

TABLE V.2

IMPACT OF PROSPECTIVE PAYMENT ON SERVICE USE, BY FOR-PROFIT OR NONPROFIT STATUS

	Treatment Group Mean[b]	Control Group Mean	Difference[b] (*p*-value)[c]
Service Use (In the Year After Admission)[a]			
Total Visits			
For-profit	71.0	87.6	-16.6***
Nonprofit	43.4	62.9	-19.5***
Subgroup difference			2.9 (0.67)
Skilled Nursing Visits			
For-profit	31.2	42.3	-11.1***
Nonprofit	19.6	27.1	-7.5***
Subgroup difference			-3.6 (0.31)
Home Health Aide Visits			
For-profit	28.8	35.4	-6.6**
Nonprofit	17.9	25.1	-7.2***
Subgroup difference			0.6 (0.91)
Physical Therapy Visits			
For-profit	8.5	7.0	1.5
Nonprofit	4.5	8.0	-3.5***
Subgroup difference			5.0*** (0.00)
Service Length (Through One Year After Admission)[a]			
Days Until Last Visit (Max =366)			
For-profit	118.6	142.0	-23.4***
Nonprofit	100.2	126.6	-26.4***
Subgroup difference			3.0 (0.79)
Probability of Last Visit By Day 120			
For-profit	67.7	58.5	9.2***
Nonprofit	72.8	64.2	8.6***
Subgroup difference			0.6 (0.94)

	For-Profit Status		**Nonprofit Status**	
Sample Size	**Patients**	**Agencies**	**Patients**	**Agencies**
Treatment Group	**5,422**	**19**	**20,962**	**20**
Control Group	**6,918**	**21**	**18,831**	**19**

SOURCE: Medicare claims data.

NOTE: Individual observations have been weighted to give agencies equal representation in the calculations.

[a] Reflects a patient's first admission to a home health agency participating in the demonstration.

[b] Regression-adjusted (through ordinary least squares and logit models) to control for preexisting differences between treatment and control agencies.

[c] The *p*-value corresponds to a test of whether the subgroup difference is statistically different from zero. It is based on standard errors inflated to account for the effects of clustering and weighting.

**p*-value < .10, two-tailed test, corresponding to a test of whether the treatment-control difference within subgroup is different from zero.

***p*-value < .05, two-tailed test, corresponding to a test of whether the treatment-control difference within subgroup is different from zero.

****p*-value < .01, two-tailed test, corresponding to a test of whether the treatment-control difference within subgroup is different from zero.

larger reduction in physical therapy visits by nonprofit agencies. Specifically, although nonprofit agencies cut the number of physical therapy visits by 3.5 per patient-year, for-profit agencies actually increased the number, by a small (and statistically insignificant) amount. This difference is difficult to explain, but it may be due to the relatively large therapy staffs of many nonprofit agencies, which afforded them a greater opportunity to make large-scale cuts in services.[3]

Because high-use agencies were much more likely than low-use agencies to be for profit, the impacts by practice pattern masked important differences between for-profit and nonprofit agencies.[4] Specifically, when we reestimated the impacts for subgroups defined by profit status and *did not* account for differences in impacts by practice pattern, we found that for-profit agencies cut about 11 more visits than did nonprofit agencies (not shown). While this difference was not statistically significant (its *p*-value is 0.12), it was quite large. Moreover, significant differences did exist in the number of home health aide visits provided and in each measure of service length. Thus, for-profit status did have an important bearing on agencies' response to prospective payment, but it was only through the greater propensity of for-profit agencies to have high-use practice patterns prior to the demonstration. For agencies with comparable predemonstration practice patterns, we find no difference between nonprofit and for-profit agencies.

Small agencies (those providing fewer than 30,000 visits during the base-year) made larger reductions in visits than did large agencies (Table V.3). Small agencies cut an average of 31 visits per patient-year, but large agencies cut only 13 visits--a marginally significant

---

[3] Based on the data from the 1996 cost reports, 14 nonprofit agencies had staffs that included five or more full-time equivalent therapists. Only five for-profit agencies had therapy staffs this large.

[4] Of the 42 high-use agencies, 26 are for profit and only 16 are nonprofit. Conversely, of the 37 low-use agencies, only 14 are for profit and 23 are nonprofit.

TABLE V.3

IMPACT OF PROSPECTIVE PAYMENT ON SERVICE USE, BY SMALL OR LARGE SIZE[a]

	Treatment Group Mean[c]	Control Group Mean	Difference[c] (*p*-value)[d]
Service Use (In the Year After Admission)[b]			
Total Visits			
Small size	70.5	101.6	-31.1***
Large size	56.4	69.4	-13.0***
Subgroup difference			-18.1* (0.10)
Skilled Nursing Visits			
Small size	33.6	44.6	-11.0*
Large size	24.0	32.7	-8.7
Subgroup difference			-2.3 (0.70)
Home Health Aide Visits			
Small size	31.2	47.2	-16.0***
Large size	22.9	26.3	-3.4**
Subgroup difference			-12.6** (0.05)
Physical Therapy Visits			
Small size	4.8	7.1	-2.3**
Large size	7.1	7.6	-0.5
Subgroup difference			-1.8 (0.22)
Service Length (Through One Year After Admission)[b]			
Days Until Last Visit (Maximum = 366)			
Small size	113.7	157.2	-43.5***
Large size	111.3	129.0	-17.7***
Subgroup difference			-25.8* (0.06)
Probability of Last Visit by Day 120			
Small size	69.5	54.0	15.5***
Large size	69.3	63.0	6.3***
Subgroup difference			9.2** (0.05)

	Small Size		**Large Size**	
Sample Size	**Patients**	**Agencies**	**Patients**	**Agencies**
Treatment Group	**2,013**	**14**	**24,371**	**25**
Control Group	**768**	**8**	**24,981**	**32**

SOURCE: Medicare claims data.

NOTE: Individual observations have been weighted to give agencies equal representation in the calculations.

[a] Large (small) agencies are defined as those that provided more (fewer) than 30,000 visits during the predemonstration year.

[b] Reflects a patient's first admission to a home health agency participating in the demonstration.

[c] Regression-adjusted (through ordinary least squares and logit models) to control for preexisting differences between treatment and control agencies.

[d] The *p*-value corresponds to a test of whether the subgroup difference is statistically different from zero. It is based on standard errors inflated to account for the effects of clustering and weighting.

**p*-value < .10, two-tailed test, corresponding to a test of whether the treatment-control difference within subgroup is different from zero.
***p*-value < .05, two-tailed test, corresponding to a test of whether the treatment-control difference within subgroup is different from zero.
****p*-value < .01, two-tailed test, corresponding to a test of whether the treatment-control difference within subgroup is different from zero.

difference of 18 visits.[5] As was the case with high-use agencies, the relatively larger reduction by small agencies resulted mostly from deeper cuts in home health aide visits. In addition, small agencies made significantly greater reductions in their length of service than did large agencies. Large agencies reduced their average length of service by 18 days, but small agencies reduced this length by 44 days. Thus, much earlier discharge appears to have driven the deeper cuts in visits that small agencies made.

Like for-profit agencies, small agencies may have cut more visits simply because their practice patterns were more often high use than low use. The most obvious evidence supporting this explanation is that small treatment agencies continued to provide about 14 visits more per patient-year than did large agencies, even after cutting more than twice the number of visits during the demonstration. This explanation, if true, raises the question of why regression models that control for impact differences by practice pattern displayed relatively greater impacts for small agencies than for large ones (but not differences by profit status). It is likely that the answer rests with our measure of predemonstration practice patterns. Given our available data for the base year, we measured agencies' practice patterns through their average (base-year) visits in the *first 120 days from admission.* However, compared with large control agencies, a relatively large share of the service use in a patient-year for most small control agencies took place after the first 120 days. Thus, some small agencies that were not identified as high use might have been identified as such had we been able to measure practice patterns over the course of a full patient-year. This possible unmeasured difference in practice patterns should have little effect on our principal findings, but it does suggest

---

[5] The impacts by agency size have limited statistical power because the sample of patients in small agencies was small. Thus, even very large differences in impacts have only marginal significance.

that small agencies (like for-profit ones) may have reduced more visits under the demonstration than did their counterparts simply because of their greater propensity to have high-use practice patterns.[6]

The responses of freestanding agencies and hospital-based agencies to the demonstration differed substantially (Table V.4); however, our sample of only nine hospital-based agencies limits the statistical significance of this difference. Freestanding agencies cut 19 visits, whereas hospital-based agencies cut 10 visits--a very large but statistically insignificant difference of 9 visits. Given this lack of significance, the difference alone does not provide sufficient evidence to conclude that prospective payment led to relatively deeper cuts by freestanding agencies. However, an examination of the provision of individual services offers additional evidence. First, hospital-based agencies made significant reductions only in skilled nursing visits, but freestanding agencies made large and significant cuts in all three service types examined (skilled nursing, home health aide, and physical therapy). Second, despite the lack of precision for the hospital-based subgroup (owing to the small sample size), the reduction in home health aide visits by freestanding agencies was statistically larger than the insignificant change made by hospital-based agencies. We thus conclude cautiously from these results that prospective payment did lead to relatively larger reductions in visits by freestanding agencies.

Like the differences by agency size, those by agency auspice may have been the result of unmeasured differences in practice patterns. As was the case with small agencies, freestanding

---

[6] In sensitivity tests (not reported), we estimated selected impacts using an additional agency-level control variable for the rate of base-year episodes continuing past 120 days, a proxy measure for preexisting practice patterns later from admission. The (main) impacts for the full sample changed only slightly with this additional variable. However, for the subgroups defined by agency size, the differences in the impacts weakened more substantially. With respect to total visits, for example, the difference in subgroup impacts fell from about 18 visits to 14 visits (with a *p*-value of 0.26), a large enough change to suggest that unobserved differences in practice patterns might explain the differences in impacts between large and small agencies.

TABLE V.4

IMPACT OF PROSPECTIVE PAYMENT ON SERVICE USE, BY HOSPITAL-BASED OR FREESTANDING

	Treatment Group Mean[b]	Control Group Mean	Difference[b] (*p*-value)[c]
Service Use (In the Year After Admission)[a]			
Total Visits			
Hospital based	44.2	54.2	-10.0
Freestanding	60.6	79.7	-19.1***
Subgroup difference			9.1 (0.22)
Skilled Nursing Visits			
Hospital based	13.1	25.2	-12.1***
Freestanding	27.8	36.8	- 9.0***
Subgroup difference			-3.1 (0.44)
Home Health Aide Visits			
Hospital based	21.4	20.4	1.0
Freestanding	24.4	32.3	-7.9***
Subgroup difference			8.9** (0.05)
Physical Therapy Visits			
Hospital based	7.4	5.7	1.7
Freestanding	6.5	7.8	-1.3* *
Subgroup difference			3.0 (0.15)
Service Length (Through One Year After Admission)[a]			
Days Until Last Visit (Maximum = 366)			
Hospital based	95.9	113.8	-17.9
Freestanding	112.5	138.3	-25.8***
Subgroup difference			7.9 (0.61)
Probability of Last Visit By Day 120			
Hospital based	76.2	68.2	8.0*
Freestanding	69.0	60.0	9.0***
Subgroup difference			-1.0 (0.87)

	Hospital Based		**Freestanding**	
Sample Size	**Patients**	**Agencies**	**Patients**	**Agencies**
Treatment Group	**868**	**3**	**25,516**	**36**
Control Group	**4,417**	**6**	**21,332**	**34**

SOURCE: Medicare claims data.

NOTE: Individual observations have been weighted to give agencies equal representation in the calculations.

[a] Reflects a patient's first admission to a home health agency participating in the demonstration.

[b] Regression-adjusted (ordinary least squares and logit models) to control for preexisting differences between treatment and control agencies.

[c] The *p*-value corresponds to a test of whether the subgroup difference is statistically different from zero. It is based on standard errors inflated to account for the effects of clustering and weighting.

**p*-value < .10, two-tailed test, corresponding to a test of whether the treatment-control difference within subgroup is different from zero.

***p*-value < .05, two-tailed test, corresponding to a test of whether the treatment-control difference within subgroup is different from zero.

****p*-value < .01, two-tailed test, corresponding to a test of whether the treatment-control difference within subgroup is different from zero.

agencies provided far more visits than did their counterparts, and much of this difference took place after 120 days, a period that our available measure of practice patterns could not capture. As a result, we cannot reject the hypothesis that the cuts in service use resulting from prospective payment would be the same for a hospital-based agency and a freestanding agency with similar practice patterns.

Our sample includes only 12 agencies with base-year costs above the limits. We nevertheless find that these agencies made significantly larger cuts in services than did agencies below their limits (Table V.5). For agencies above the limits, prospective payment led to reductions of about 31 visits during the patient-year--a reduction roughly twice the size made by agencies below the limits. The larger impact for agencies above the limits is at least partly due to differences in the behavior of control group members. Specifically, between the base year and first demonstration year, control agencies above the limits substantially increased their average visits per patient, while agencies below the limits made little change.[7] This difference is not surprising; as noted, a typical control agency above the limits has a particularly strong incentive to increase visits to reduce its per visit costs.

C. IMPACTS FOR PATIENT SUBGROUPS

We find no clear evidence that impacts differed for patients who could or could not take oral medications independently (Table V.6). For each of the six outcomes shown, the reductions in service use and length were actually larger for patients who could not take these medications independently, although they did not differ by a statistically significant margin and were similar as a proportion. For some outcomes, particularly total visits and home health aide visits, the differences

---

[7] Data not shown. Due to data limitations in the base year, this comparison was limited to service use during the first 120 days after a patient admission.

TABLE V.5

IMPACT OF PROSPECTIVE PAYMENT ON SERVICE USE, BY UNDER OR OVER (BASE-YEAR) COST LIMITS

	Treatment Group Mean[b]	Control Group Mean[b]	Difference[b] (*p*-value)[c]
Service Use (In the Year After Admission)[a]			
Total Visits			
Under cost limit	59.4	75.1	-15.7***
Over cost limit	37.8	68.9	-31.1***
Subgroup difference			15.4* (0.06)
Skilled Nursing Visits			
Under cost limit	25.9	34.2	-8.3***
Over cost limit	19.4	34.5	-15.1***
Subgroup difference			6.8 (0.16)
Home Health Aide Visits			
Under cost limit	24.5	30.2	-5.8**
Over cost limit	12.8	25.7	-12.9***
Subgroup difference			7.1 (0.16)
Physical Therapy Visits			
Under cost limit	6.8	7.5	-0.7
Over cost limit	4.6	7.1	-2.5
Subgroup difference			1.8 (0.38)
Service Length (Through One Year After Admission)[a]			
Days Until Last Visit (Maximum = 366)			
Under cost limit	116.3	136.8	-20.5***
Over cost limit	83.2	132.7	-49.5***
Subgroup difference			29.0*** (0.01)
Probability of Last Visit by Date 120			
Under cost limit	68.0	60.7	7.3***
Over cost limit	78.1	60.1	18.0***
Subgroup difference			-10.7*** (0.00)

	Under Cost Limit		**Over Cost Limit**	
Sample Size	**Patients**	**Agencies**	**Patients**	**Agencies**
Treatment Group	**24,645**	**33**	**1,739**	**6**
Control Group	**21,333**	**34**	**4,416**	**6**

SOURCE: Medicare claims data.

NOTE: Individual observations have been weighted to give agencies equal representation in the calculations.

[a] Reflects a patient's first admission to a home health agency participating in the demonstration.

[b] Regression-adjusted (through ordinary least squares and logit models) to control for preexisting differences between treatment and control agencies.

[c] The *p*-value corresponds to a test of whether the subgroup difference is statistically different from zero. It is based on standard errors inflated to account for the effects of clustering and weighting.

**p*-value < .10, two-tailed test, corresponding to a test of whether the treatment-control difference within subgroup is different from zero.

***p*-value < .05, two-tailed test, corresponding to a test of whether the treatment-control difference within subgroup is different from zero.

****p*-value < .01, two-tailed test, corresponding to a test of whether the treatment-control difference within subgroup is different from zero.

TABLE V.6

IMPACT OF PROSPECTIVE PAYMENT ON SERVICE USE. BY WHETHER PATIENT CAN TAKE ORAL MEDICATIONS INDEPENDENTLY[a]

	Treatment Group Mean[c]	Control Group Mean	Difference[c] (*p*-value)[d]
Service Use (In the Year After Admission)[b]			
Total Visits			
Can take oral medications independently	41.6	54.3	-12.7**
Cannot take oral medications independently	60.6	80.4	-19.8***
Subgroup difference			7.1 (0.25)
Skilled Nursing Visits			
Can take oral medications independently	22.0	28.6	-6.6**
Cannot take oral medications independently	25.9	34.4	-8.5***
Subgroup difference			1.9 (0.50)
Home Health Aide Visits			
Can take oral medications independently	12.7	17.8	-5.1
Cannot take oral medications independently	25.8	35.4	-9.6**
Subgroup difference			4.5 (0.33)
Physical Therapy Visits			
Can take oral medications independently	5.8	6.3	-0.5
Cannot take oral medications independently	6.8	7.5	-0.7
Subgroup difference			0.2 (0.80)
Service Length (Through One Year After Admission)[b]			
Days Until Last Visit (Maximum = 366)			
Can take oral medications independently	93.8	112.6	-18.8***
Cannot take oral medications independently	113.2	136.6	-23.4***
Subgroup difference			4.6 (0.53)
Probability of Last Visit by Day 120			
Can take oral medications independently	75.6	69.2	6.4***
Cannot take oral medications independently	68.9	60.8	8.1***
Subgroup difference			-1.7 (0.66)

	Can Take Oral Medications Independently		**Cannot Take Oral Medications Independently**	
Sample Size	**Patients**	**Agencies**	**Patients**	**Agencies**
Treatment Group	**5,696**	**35**	**6,170**	**35**
Control Group	**2,858**	**35**	**3,500**	**35**

SOURCE: Medicare claims data.

NOTE: Individual observations have been weighted to give agencies equal representation in the calculations.

[a] A patient is able to take oral medications independently if quality assurance data collected at admission indicates this is the case. Otherwise, the patient is defined as not able to take oral medications independently.

[b] Reflects a patient's first admission to a home health agency participating in the demonstration.

[c] Regression-adjusted (through ordinary least squares and logit models) to control for preexisting differences between treatment and control agencies.

[d] The *p*-value corresponds to a test of whether the subgroup difference is statistically different from zero. It is based on standard errors inflated to account for the effects of clustering and weighting.

**p*-value < .10, two-tailed test, corresponding to a test of whether the treatment-control difference within subgroup is different from zero.
***p*-value < .05, two-tailed test, corresponding to a test of whether the treatment-control difference within subgroup is different from zero.
****p*-value < .01, two-tailed test, corresponding to a test of whether the treatment-control difference within subgroup is different from zero.

in impacts were quite large, suggesting that we may simply lack the statistical power to detect actual differences. This conclusion is difficult to embrace, however, because the estimated differences in impacts were actually in the opposite direction than we expected. Nevertheless, the fact that these declines were even similar to one another suggests that treatment agencies were not more reluctant to reduce services for patients who could not take oral medications independently than for those who could. Because patients with the limitation might be expected to be relatively more adversely affected by reductions in service use, our future analyses (particularly of patient outcomes and other Medicare service use) should examine this subgroup in some detail.

We find no evidence that agencies made differential reductions in services for patients with and without other sources of care (Table V.7). The estimated reductions in visits and service length were similar for the two subgroups shown in Table V.7, and no significant differences existed between any of the outcomes examined. (Indeed, the similarity of the control group means for the two subgroups suggests little difference in the service demands of the two patient types.) The absence of significant differences between these subgroups is surprising, as a number of treatment agencies noted during site visits that they had increased their reliance on family members and other informal caregivers. One possible explanation for this result, as well as the similarity in the control group means, is the measure of caregiver availability, which is based on information at admission, and which therefore does not account for possible changes in caregiver availability that may mitigate observed differences between subgroups. Assuming such changes were relatively infrequent, however, our findings suggest that the efforts by treatment agencies to rely more heavily on informal caregivers were not dramatic enough to generate substantive reductions in services. In addition, they suggest that the service needs of the two patient types may not differ greatly on average.

TABLE V.7

IMPACT OF PROSPECTIVE PAYMENT ON SERVICE USE, BY WHETHER OTHER CAREGIVERS AVAILABLE[a]

	Treatment Group Mean[c]	Control Group Mean	Difference[c] (*p*-value)[d]
Service Use (In the Year After Admission)[b]			
Total Visits			
No other caregiver	53.4	70.8	-17.4***
Other caregiver	51.9	68.5	-16.6***
Subgroup difference			-0.8 (0.86)
Skilled Nursing Visits			
No other caregiver	25.3	33.2	-7.9***
Other caregiver	23.4	31.1	-7.7***
Subgroup difference			-0.2 (0.93)
Home Health Aide Visits			
No other caregiver	20.3	29.0	-8.7***
Other caregiver	20.3	27.4	-7.1**
Subgroup difference			-1.6 (0.60)
Physical Therapy Visits			
No other caregiver	6.1	6.2	-0.1
Other caregiver	6.5	7.5	-1.0
Subgroup difference			0.9 (0.23)
Service Length (Through One Year After Admission)[b]			
Days Until Last Visit (Max = 366)			
No other caregiver	104.6	132.8	-28.2***
Some other caregiver	104.9	122.3	-17.4***
Subgroup difference			-10.8 (0.13)
Probability of Last Visit By Day 120			
No other caregiver	71.5	61.6	9.9***
Other caregiver	72.0	66.1	5.9**
Subgroup difference			4.0 (0.13)

	No Other Caregiver		**Other Caregiver**	
Sample Size	**Patients**	**Agencies**	**Patients**	**Agencies**
Treatment Group	**5,178**	**35**	**6,688**	**35**
Control Group	**2,495**	**35**	**3,863**	**35**

SOURCE: Medicare claims data.

NOTE: Individual observations have been weighted to give agencies equal representation in the calculations.

[a] A patient is defined as having another caregiver available if data collected at admission indicates any live-in informal or paid care is available to the patient, or if the patient resides in an assisted-living facility. Otherwise, the patient is defined as having no informal caregiver available.

[b] Reflects a patient's first admission to a home health agency participating in the demonstration.

[c] Regression-adjusted (through ordinary least squares and logit models) to control for preexisting differences between treatment and control agencies.

[d] The *p*-value corresponds to a test of whether the difference is statistically subgroup different from zero. It is based on standard errors inflated to account for the effects of clustering and weighting.

**p*-value < .10, two-tailed test, corresponding to a test of whether the treatment-control difference within subgroup is different from zero.

***p*-value < .05, two-tailed test, corresponding to a test of whether the treatment-control difference within subgroup is different from zero.

****p*-value < .01, two-tailed test, corresponding to a test of whether the treatment-control difference within subgroup is different from zero.

TABLE V.8

IMPACT OF PROSPECTIVE PAYMENT ON SERVICE USE, BY HIGH OR LOW EXPECTED COSTS[a]

	Treatment Group Mean[c]	Control Group Mean	Difference[c] (*p*-value)[d]
Service Use (In the Year After Admission)[b]			
Total Visits			
High expected costs	84.1	109.0	-24.9***
Low expected costs	35.6	47.7	-12.1***
Subgroup difference			-12.8* (0.10)
Skilled Nursing Visits			
High expected costs	34.0	44.8	-10.8
Low expected costs	18.8	24.9	-6.1
Subgroup difference			-4.7 (0.21)
Home Health Aide Visits			
High expected costs	39.9	52.8	-12.9
Low expected costs	9.8	14.4	-4.6
Subgroup difference			-8.3 (0.15)
Physical Therapy Visits			
High expected costs	7.7	8.1	-0.4
Low expected costs	5.7	6.4	-0.7
Subgroup difference			0.3 (0.78)
Service Length (Through One Year After Admission)[b]			
Days Until Last Service (Maximum = 365)			
High expected costs	120.1	146.9	-26.8
Low expected costs	96.4	115.3	-18.9
Subgroup difference			-7.9 (0.24)
Probability of Last Visit By Day 120			
High expected costs	66.8	57.3	9.5
Low expected costs	74.6	68.2	6.4
Subgroup difference			3.1 (0.40)

	High Expected Costs		**Low Expected Costs**	
Sample Size	**Patients**	**Agencies**	**Patients**	**Agencies**
Treatment Group	**3,555**	**35**	**8,311**	**35**
Control Group	**1,919**	**35**	**4,439**	**35**

SOURCE: Medicare claims data.

NOTE: Individual observations have been weighted to give agencies equal representation in the calculations

[a] A patient is defined as having high expected costs if his/her predicted reimbursement for home health services in the year after admission is in the top quartile for all patients in the sample. Otherwise, the patient is defined as having low expected costs.

[b] Reflects a given patient's first admission to a home health agency participating in the demonstration.

[c] Regression-adjusted (through ordinary least squares and logit models) to control for preexisting differences between treatment and control agencies.

[a] The *p*-value corresponds to a test of whether the subgroup difference is statistically different from zero. It is based on standard errors inflated to account for the effects of clustering and weighting.

**p*-value < .10, two-tailed test, corresponding to a test of whether the treatment-control difference within subgroup is different from zero.

***p*-value < .05, two-tailed test, corresponding to a test of whether the treatment-control difference within subgroup is different from zero.

****p*-value < .01, two-tailed test, corresponding to a test of whether the treatment-control difference within subgroup is different from zero.

Consistent with our expectation, we find some evidence to suggest that treatment agencies made significantly larger reductions in visits for patients with relatively high expected costs (Table V.8). Overall, prospective payment significantly reduced average visits for patients in the low-cost subgroup from about 48 to 36 visits (25 percent), while it cut average visits in the high-cost subgroup from about 110 to 85 visits (23 percent). The impact among high-cost patients was thus nearly twice the magnitude, though only marginally significant (p=.10) and similar as a proportion. The absolute difference in impacts suggests that treatment agencies generated far more savings per high-cost patient than they did per low-cost patient. However, the similarity in proportional effects suggests that they generated about the same rate of savings from the two patient types. None of the remaining outcomes differed significantly; however, each was generally large and consistent with the evidence of greater overall cuts for high-cost patients.

D. SUMMARY

Agencies in all subgroup categories that we examined responded to prospective payment by making significant reductions in services for different types of patients. The subgroup findings thus confirm the robustness of our main findings, as overall program impacts are not driven by very large effects for a subset of agencies.

Although we find consistently substantial program impacts across all different types of agencies, the size of the impacts varied significantly between certain subgroups. With respect to agency subgroups, we find:

- Very strong evidence that agencies with high-use practice patterns prior to the demonstration made deeper cuts in the number of and duration of services than did agencies with low-use practice patterns

- Strong evidence that for-profit agencies cut more visits than did nonprofit ones. This difference occurred entirely because for-profit agencies were more likely than nonprofit agencies to have high-use practice patterns.

- Strong evidence that relatively small agencies made larger reductions in service use and duration than did large agencies. Some evidence suggests that this difference was due to the relatively greater propensity of small agencies to have high-use practice patterns.

- Compelling evidence that freestanding agencies cut more visits (particularly home health aide visits) than did hospital-based agencies. Some evidence again suggests that this difference was due to the relatively greater propensity of freestanding to have high-use practice patterns.

- Compelling evidence that agencies above the cost limits cut more visits and shortened service length more than did agencies below the limits

With respect to patient subgroups, we find:

- Compelling evidence that agencies cut more visits for patients expected to require a disproportionate share of services. However, the proportionate reduction for patients who did not fall into this category was very similar.

- Little evidence that agencies reduced services differently for patients who could or could not take oral medications independently, or for patients who did or did not have other caregivers available.

For the subgroups experiencing the largest reduction in services, it will be particularly important for future reports to examine whether these reductions led to more adverse effects on patient functioning and satisfaction. Similarly, it will be important to focus on patients who cannot take oral medications independently, because these patients experienced surprisingly large declines in service use, equal to or perhaps even larger than those of other patients.

VI. COMPARISON BETWEEN DEMONSTRATION YEARS ONE AND TWO

To the extent possible, it is important to understand how impacts may have changed over the course of the demonstration in order to more accurately measure the potential savings from prospective payment. In the analyses presented in Chapters IV and V, we focused on patients admitted to agencies during the first demonstration year only because we have a full 12 months of follow-up data on too few patients admitted in year two.[1] In this chapter, we therefore compare impacts on key outcomes between the first and second demonstration years over two shorter periods--the first four months after admission, and the first eight months after admission.

Each of these analyses offers different advantages in detecting changes in program impacts. If differences between years existed but were fairly modest, we might have a greater chance of detecting them in the four-month than in the eight-month analysis because we observe admissions later into the second year and have a larger overall sample. However, the eight-month analysis might enable us to detect differences that developed slowly after admission, or that grew considerably only after the first 120 days.

A. EXPECTED IMPACTS

Several factors suggest that the demonstration impacts should increase in year two, although this outcome is far from certain. Certainly most initiatives to reduce visits take time to implement, suggesting that the effects of prospective payment should strengthen as the demonstration continues.

---

[1] For 46 agencies that started the demonstration on January 1, 1996, we have almost no admissions after the first demonstration year for which we have 12 months (one "patient-year") of complete data. For the remaining 33 agencies, which started in 1995, we have some year-two admissions that are complete through 12 months, but we excluded them from the main analysis in order to have a consistently defined sample across agencies. For a further discussion of this issue, see Chapter III.

Indeed, a number of treatment agencies reported that they were slow to respond to demonstration incentives because they had to first address billing, cash flow, and other problems associated with the change to prospective payment. In addition, some treatment agencies reported that their staff initially resisted reducing the frequency of visits or pursuing faster discharge. Some agencies also reported that the demonstration procedures themselves, particularly completing the quality assurance data, delayed implementation of planned initiatives.

Other factors suggest that impacts in demonstration year two might have remained unchanged or might have weakened. The very large year one impacts indicate that a number of treatment agencies were able to implement efforts to cut visits not only successfully but rapidly. Many of these agencies that might have been able to make additional cuts during the second demonstration year might also have been concerned that continuing reductions in services could compromise patient outcomes. Agencies might also have found that initial gains in net revenue from reducing visits early in the demonstration led to a sharp increase in cost per visit, as volume over which to distribute overhead was lost.[2] Reports from some treatment agency staff indicating surprise at how dramatically their cost per visit rose during the first demonstration year support this explanation.[3] Some agencies also reported an overall decrease in patient referrals resulting from greater competition, which could increase the potential risk of additional loss of volume from cutting visits.

---

[2] Depending on the size of any increase in cost per visit and several other factors (for example, the number of outlier visits provided, the types of services cut, and the distribution of service reductions between the at-risk and outlier periods), agencies that significantly reduced visits during the first demonstration year could have experienced financial losses. The impacts of the demonstration on costs and profits/losses will be examined in a future report.

[3] Cost reports for year one confirmed the comments of staff from several treatment agencies that their cost per visit had risen sharply.

Important market or "environmental" factors might also have affected the size of the impacts between the two demonstration years, most likely weakening the relative impacts in year two. For example, the ongoing federal program against fraud and abuse (Operation Restore Trust) has caused many home health agencies to more carefully justify their care and has probably reduced average visits per patient in areas under investigation. Likewise, the growing influence of managed care in the home health market may be driving visits down for Medicare patients as agencies implement programs to provide quality service under a restricted plan of care. Because treatment agencies are already cutting visits in response to prospective payment, we expect that these environmental factors would affect control agencies more strongly than they would treatment agencies. In general, this differential effect would tend to reduce the impact of the demonstration in the second year, because (all else equal) the difference in visits between control and treatment groups would be smaller in that year.

B. DIFFERENCES AT FOUR AND EIGHT MONTHS FROM ADMISSION

Based on the four-month (at-risk period) analysis of service provision, we find no evidence that the impact of prospective payment on the average visits provided to patients changed between the first and second years of the demonstration (Table VI.I). In each month of the at-risk period, treatment agencies provided significantly fewer visits on average than did control agencies, and this difference remained very similar between years. By the end of the at-risk period, prospective payment had eliminated nearly eight visits per patient in both years.

TABLE VI.1

COMPARISON OF YEAR ONE AND YEAR TWO IMPACTS ON TOTAL VISITS PROVIDED IN THE FIRST 120 DAYS AFTER ADMISSION[a]

	Year One			Year Two			
Total Visits from Admission Through:	Treatment Group Mean[b]	Control Group Mean	Difference[b] (*p*-value)[c]	Treatment Group Mean[b]	Control Group Mean	Difference[b] (*p*-value)[c]	Change in Impacts, Year 2 - Year 1[b] (*p*-value)[c]
At-Risk Period							
30 Days	17.8	19.4	-1.6*** (0.01)	16.7	18.3	-1.6** (0.02)	0.0 (0.92)
60 Days	27.1	30.9	-3.8*** (0.00)	24.8	28.8	-4.0*** (0.00)	-0.2 (0.80)
90 Days	32.8	38.5	-5.7*** (0.00)	29.8	35.7	-5.9*** (0.00)	-0.2 (0..80)
120 Days	37.3	44.8	-7.5*** (0.00)	33.7	41.4	-7.7*** (0.00)	-0.2 (0.86)
Sample Size:							
Patients	**26,384**	**25,749**	--	**15,365**	**17,553**	--	--
Agencies	**39**	**40**		**39**	**40**		

SOURCE: Medicare claims data.

NOTE: Individual observations (patient records) have been weighted to give agencies equal representation in the calculations.

[a] Reflects a patient's first admission to a home health agency participating in the demonstration.

[b] Regression-adjusted (through ordinary least squares models) to control for preexisting differences between treatment and control agencies.

[c] The *p*-value corresponds to a test of whether the treatment-control difference (impact) is statistically different from zero. It is based on standard errors inflated to account for the effects of clustering and weighting.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.
***Significantly different from zero at the .01 level, two-tailed test.

Despite the similarity of the impacts between years, both treatment and control agencies actually made significant--but offsetting--reductions in visits during the second year of the demonstration.[4] Average visits provided by treatment agencies fell from 37.3 in year one to 33.7 in year two--a significant decrease of 3.6 visits ($p < .01$, not shown), or close to half the demonstration impact in year one (7.5 visits). Average visits provided by control agencies also fell significantly, from 44.8 in year one to 41.4 in year two ($p = .01$, not shown), almost the same decrease as in treatment agencies. The net result was a demonstration impact in year two that was almost exactly the size of the impact in year one.

Interestingly, the decrease among control agencies between the first and second years was far larger than that between the predemonstration and first years; this pattern did not hold up for treatment agencies. During the base year (the year before the demonstration), control agencies provided an average of 45.7 visits, compared with 44.8 in the first year of the demonstration.[5] This modest drop (0.9 visits) was thus only a small fraction of the 3.4 visit decrease between the first two demonstration years. In contrast, treatment agencies' average visits fell nearly nine visits between the base year and first demonstration year, more than twice the size of the decrease between demonstration years one and two.

---

[4] To recover unbiased impact estimates, our sample includes only a patient's initial admission to an agency during the demonstration (see Chapter III for details). This leads to slightly different sample construction between the two demonstration years, because our year-two sample excludes patients admitted to a demonstration agency during the preceding year, whereas the year-one sample does not. These sample differences inflate the reduction in visits by both treatment and control agencies shown in Table VI.1 (and in later tables), because the average visits for a patient admitted in both years one and two was higher than for a patient admitted in year two only. However, the inflation was small and does not affect the significance of any changes between demonstration years.

[5] These changes were shown in Table IV.12.

The reduction in visits by control agencies between demonstration years one and two was largely centered in skilled nursing and home health aide services, and it was large enough in the case of skilled nursing visits to reduce the demonstration impact between the two years (Table VI.2). During the first year, prospective payment significantly reduced skilled nursing visits in the first 120 days by an average of 4 visits; in year two, the impact of prospective payment remained significant but fell to 2.7 visits. The reason for this decline (about one-third of the year one impact) is that control agencies eliminated nearly two skilled nursing visits per patient in year two, a reduction that more than offset an additional decline among treatment agencies of 0.5 skilled nursing visits. The pattern for home health aide visits was the reverse; treatment and control agencies respectively cut about 2.6 and 1.8 home health aide visits between demonstration years, resulting in a small (but statistically insignificant) increase in the year-two impact. Changes between the years for the other services were small and similar for both groups.

Estimates of demonstration impacts at eight months from admission support the findings from the four-month analysis (Table VI.3). At eight months, prospective payment reduced the total visits per patient significantly by about 13 visits in year one and by about 12 visits in year two. These impacts did not differ statistically; however, large declines were observed by both treatment and control agencies between the two demonstration years.[6]

The reduction in the demonstration's impact on skilled nursing visits from year one to year two during the at-risk period was larger and even more significant over the eight-month period, again

---

[6] Unlike the findings at four month, declines by treatment and control agencies between the two demonstration years were not statistically significant (*p*-values = 0.25 and 0.24, respectively). This is explained by a smaller sample size and a much larger design effect for the eight month sample, which resulted in a relative loss of precision.

TABLE VI.2

COMPARISON OF YEAR ONE AND YEAR TWO IMPACTS ON THE NUMBER OF VISITS PROVIDED IN THE FIRST 120 DAYS (AT-RISK PERIOD) AFTER ADMISSION, BY SERVICE TYPE[a]

	Year One			Year Two			
Number of Visits	Treatment Group Mean[b]	Control Group Mean	Difference[b] (*p*-value)[c]	Treatment Group Mean[b]	Control Group Mean	Difference[b] (*p*-value)[c]	Change in Impacts, Year 2 – Year 1[b] (*p*-value)[c]
All	37.3	44.8	-7.5*** (0.00)	33.7	41.4	-7.7*** (0.00)	-0.2 (0.87)
Skilled Nursing	17.6	21.6	-4.0*** (0.00)	17.1	19.8	-2.7*** (0.02)	1.3* (0.06)
Home Health Aide	12.8	15.5	-2.7*** (0.00)	10.2	14.1	-3.9*** (0.00)	-1.2 (0.16)
Physical Therapy	5.2	5.6	-0.4 (0.45)	4.9	5.5	-0.6 (0.27)	-0.2 (0.41)
Occupational Therapy	0.8	1.0	-0.2 (0.20)	0.7	1.0	-0.3** (0.05)	-0.1 (0.27)
Speech Therapy	0.3	0.3	0.0 (0.48)	0.3	0.3	0.0 (0.82)	0.0 (0.93)
Medical Social Worker	0.5	0.8	-0.3*** (0.01)	0.4	0.7	-0.3*** (0.01)	0.0 (0.82)
Sample Size:							
Patients	**26,384**	**25,749**	--	**15,365**	**17,553**	--	--
Agencies	**39**	**40**		**39**	**40**		

SOURCE: Medicare claims data.

NOTE: Individual observations (patient records) have been weighted to give agencies equal representation in the calculations.

[a] Reflects a patient's first admission to a home health agency participating in the demonstration.

[b] Regression-adjusted (through ordinary least squares models) to control for preexisting differences between treatment and control agencies.

[c] The *p*-value corresponds to a test of whether the treatment-control difference (impact) is statistically different from zero. It is based on standard errors inflated to account for the effects of clustering and weighting.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.
***Significantly different from zero at the .01 level, two-tailed test.

TABLE VI.3

COMPARISON OF YEAR ONE AND YEAR TWO IMPACTS ON TOTAL VISITS PROVIDED IN THE FIRST EIGHT MONTHS AFTER ADMISSION[a]

	Year One			Year Two			
Total Visits from Admission Through:	Treatment Group Mean[b]	Control Group Mean	Difference (*p*-value)[c]	Treatment Group Mean[b]	Control Group Mean	Difference[b] (*p*-value)[c]	Change in Impacts, Year 2 − Year 1[b] (*p*-value)[c]
At-Risk Period							
30 Days	17.8	19.4	-1.6*** (0.01)	17.8	18.2	-1.4** (0.05)	0.2 (.74)
60 Days	27.1	30.9	-3.8*** (0.00)	25.3	28.9	-3.6*** (0.01)	0.2 (.84)
90 Days	32.8	38.5	-5.7*** (0.00)	30.3	35.7	-5.4*** (0.00)	0.3 (.75)
120 Days	37.3	44.8	-7.5*** (0.00)	34.4	41.5	-7.1*** (0.00)	0.4 (.80)
Post-At-Risk Period							
150 Days	40.6	50.0	-9.4*** (0.00)	37.4	46.0	-8.6*** (0.00)	0.8 (0.65)
180 Days	43.6	54.6	-11.0*** (0.00)	40.7	50.1	-9.7*** (0.00)	1.3 (0.53)
210 Days	46.3	58.7	-12.4*** (0.00)	42.8	53.5	-10.7*** (0.00)	1.7 (0.50)
240 Days	49.3	62.6	-13.3*** (0.00)	44.8	56.6	-11.8*** (0.00)	1.5 (0.49)
Sample Size:							
Patients	**26,384**	**25,749**	--	**9,556**	**12,388**	--	--
Agencies	**39**	**40**		**39**	**40**		

SOURCE: Medicare claims data

NOTE: Individual observations (patient records) have been weighted to give agencies equal representation in the calculations.

[a] Reflects a patient's first admission to a home health agency participating in the demonstration.

[b] Regression-adjusted (through ordinary least squares models) to control for preexisting differences between treatment and control agencies.

[c] The *p*-value corresponds to a test of whether the treatment-control difference (impact) is statistically different from zero. It is based on standard errors inflated to account for the effects of clustering and weighting.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.
***Significantly different from zero at the .01 level, two-tailed test.

TABLE VI.4

COMPARISON OF YEAR ONE AND YEAR TWO IMPACTS ON NUMBER OF VISITS PROVIDED IN THE FIRST EIGHT MONTHS AFTER ADMISSION, BY SERVICE TYPE[a]

	Year One			Year Two			
Number of Visits	Treatment Group Mean[b]	Control Group Mean	Difference[b] (*p*-value)[c]	Treatment Group Mean[b]	Control Group Mean	Difference[b] (*p*-value)[c]	Change in Impacts, Year 2-Year 1[b] (*p*-value)[c]
All	48.9	62.6	-13.7*** (0.00)	44.8	56.6	-11.8*** (0.00)	1.9 (0.49)
Skilled Nursing	22.5	29.4	-6.9*** (0.00)	22.5	25.9	-3.4*** (0.04)	3.5 (0.01)
Home Health Aide	18.5	23.8	-5.3*** (0.00)	15.3	21.7	-6.4*** (0.01)	-1.1 (0.54)
Physical Therapy	6.0	6.8	-0.8 (0.18)	5.4	6.7	-1.3** (0.05)	-0.5 (0.20)
Occupational Therapy	0.9	1.2	-0.3 (0.11)	0.3	1.1	-0.8** (0.04)	-0.5 (0.62)
Speech Therapy	0.4	0.4	0.0 (0.75)	0.3	0.3	0.0 (0.93)	0.0 (0.79)
Medical Social Worker	0.6	1.0	-0.4*** (0.01)	0.6	0.9	-0.3*** (0.02)	0.1 (0.30)
Sample Size:							
Patients	**26,384**	**25,749**	--	**9,556**	**12,388**	--	--
Agencies	**39**	**40**		**39**	**40**		

SOURCE: Medicare claims data.

NOTE: Individual observations (patient records) have been weighted to give agencies equal representation in the calculations.

[a] Reflects a patient's first admission to a home health agency participating in the demonstration.

[b] Regression-adjusted (through ordinary least squares models) to control for preexisting differences between treatment and control agencies.

[c] The *p*-value corresponds to a test of whether the treatment-control difference (impact) is statistically different from zero. It is based on standard errors inflated to account for the effects of clustering and weighting.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.
***Significantly different from zero at the .01 level, two-tailed test.

because of substantial decreases in the number of visits by control agencies (Table VI.4). For admissions during the first year, prospective payment eliminated an average of about seven skilled nursing visits in the first eight months. For admissions in year two, however, the impact dropped by half, to only 3.4 visits. This sharp decline is entirely explained by a reduction among control agencies of 3.5 skilled nursing visits per patient in the second demonstration year. Treatment agencies made slightly larger reductions in the number of home health aide and physical therapy visits than did control agencies, although the differences were not significant.

Consistent with the similar decreases in service provision by treatment and control agencies, agencies in both groups also reduced their length of service by nearly equal amounts (Table VI.5).[7] In each month after admission, the probability of a patient receiving a last visit from either a treatment agency or a control agency increased between demonstration years. By eight months (240 days) from admission, for example, about 87 percent of patients admitted to treatment agencies in year one had received their last visit, as did more than 90 percent in year two. Similarly, control agencies increased their rate of last visits from about 79 percent in year one to about 83 percent in year two. Therefore, the resulting treatment-control differences in the two years (about seven percentage points) were large but almost identical.

The large service decreases among treatment agencies and control agencies suggest that a number of factors influenced both groups' behaviors during the demonstration. Both groups reported an increased presence of managed care during the demonstration, which may have led all the agencies to adopt a managed-care-like approach to service provision in order to compete for

---

[7] Following the methodology for the main analysis, the number of days until the last visit (service length) is given by the time from admission until the date of discharge from the last episode begun within the first 240 days. For patients who have not been discharged by day 240, the length of service is censored at a value of 241.

TABLE VI.5

COMPARISON OF YEAR ONE AND YEAR TWO IMPACTS ON THE TIME FROM ADMISSION TO LAST SERVICE[a]

	Year One			Year Two			
Last Service Provided by:	Treatment Group Percentage[b]	Control Group Percentage	Difference[b] (*p*-value)[c]	Treatment Group Percentage[b]	Control Group Percentage	Difference[b] (*p*-value)[c]	Change in Impacts, Year 2 – Year 1[b] (*p*-value)[c]
At-Risk Period							
30 Days	33.8	26.5	7.3*** (0.00)	37.1	29.0	8.1*** (0.00)	0.8 (0.63)
60 Days	58.5	49.1	9.4*** (0.00)	62.6	52.6	10.0*** (0.00)	0.6 (0.76)
90 Days	65.6	57.2	8.4*** (0.00)	69.6	60.7	8.9*** (0.00)	0.5 (0.83)
120 Days	73.3	64.3	9.0*** (0.00)	77.4	69.6	7.8*** (0.00)	-1.2 (0.47)
Post-at-Risk Period							
150 Days	77.1	69.0	8.1*** (0.00)	81.6	73.6	8.0*** (0.00)	-0.1 (0.90)
180 Days	80.9	73.1	7.8*** (0.00)	86.0	77.9	8.1*** (0.00)	0.3 (0.87)
240 Days	86.5	79.4	7.1*** (0.00)	91.8	83.4	8.4*** (0.00)	1.3 (0.55)
Sample Size							
Patients	**26,384**	**25,749**	--	**9,556**	**12,388**	--	--
Agencies	**39**	**40**		**39**	**40**		

SOURCE: Medicare claims data.

NOTE: Individual observations (patient records) have been weighted to give agencies equal representation in the calculations.

[a] Reflects a patient's first admission to a home health agency participating in the demonstration.

[b] Regression-adjusted (through ordinary least squares models) to control for preexisting differences between treatment and control agencies.

[c] The *p*-value corresponds to a test of whether the treatment-control difference (impact) is statistically different from zero. It is based on standard errors inflated to account for the effects of clustering and weighting.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.
***Significantly different from zero at the .01 level, two-tailed test.

contracts. In addition, growing concern about ORT investigations might have led some agencies to reduce services when they would not have otherwise done so. In addition, the ORT program might have induced referring physicians to reduce the amount of care they ordered, again lowering the number of visits provided by any affected agencies. Increased market competition may have also been a factor, particularly as nondemonstration agencies under cost-based reimbursement had an incentive to "pursue" patients requiring extensive care.

Other factors may have affected the two agency groups separately. Obviously, treatment agencies might have reduced visits largely because of the ongoing incentives provided by the demonstration. Meanwhile, control agencies might have begun seeking ways to reduce visits in response to an expectation that a form of capitated payment system would soon be mandated in the home health industry.[8] Indeed, the fact that these agencies sought to participate in the demonstration suggests that they were concerned about future changes in their system of payment.

C. SUMMARY

We find strong evidence that prospective payment had a large impact on service during both the first and second year of the demonstration. For patients admitted during the first year of the demonstration, treatment agencies provided significantly fewer visits than control agencies in response to prospective payment. Treatment agencies then made further reductions in year two; however, control agencies reduced service provision by about a nearly equal amount. Thus, the year-

---

[8] The 1997 Balanced Budget Act, which mandated a prospective payment system by 1999 and a new interim payment system until then, certainly would have had a profound effect on the expectations of control agencies. Although all the admissions in our sample occurred before passage of the act, a portion of the subsequent service period for some patients admitted in demonstration year two occurred after its passage.

one treatment-control difference was maintained in year two of the demonstration, but it was not enlarged.

These changes over the course of the demonstration, particularly by the control agencies, make it considerably more difficult to interpret how prospective payment would affect agencies' service provision in the long run, when visits by both treatment and control agencies would have stabilized. Suppose, for example, that the reductions observed among treatment agencies by year two reflected a maximum response to the incentives presented by prospective payment.[9] If control agencies were to continue to reduce visits in response to market changes, such as pressure from managed care, our estimated demonstration impacts might overstate the future savings from a system of prospective payment relative to a system of cost-based reimbursement. On the other hand, if control agencies had made reductions largely in anticipation of future changes to their payment system, or if treatment agencies continued to make further cuts in services in response to the demonstration incentives, then our impact estimates might understate the effects of prospective payment.

Despite these concerns, our findings clearly produce one very important conclusion: prospective payment, as defined by the demonstration, is a highly effective tool for reducing service provision by home health agencies. Under the demonstration, an overwhelming number of treatment agencies dramatically reduced visits per patient during the first year and continued to do so into the second. Although the reduction in year two could not be attributed to prospective payment (because of similar reductions by control agencies), they underscore the substantial opportunity that exists to lower the use of home health services.

---

[9] That is, treatment agencies would make no additional cuts in visits, despite likely financial incentives. Agencies might adopt this attitude because of concern that patient outcomes would be adversely affected by further service reductions.

VII. CONCLUSIONS

Based on a thorough analysis of data from the first two years of the Per-Episode Home Health Prospective Payment Demonstration, we find very strong and consistent evidence that prospective payment generated substantial, broad-based reductions in the number and duration of services provided to patients. The consistency and robustness of the findings suggest that they can serve as a useful guide to policymakers seeking to implement some form of prospective payment system as mandated by the Balanced Budget Act of 1997.

A. KEY FINDINGS

Prospective payment (as defined by the rules of the demonstration) significantly reduced the average number of visits provided to home health patients during the first 12 months after admission. For patients admitted in year one of the demonstration, those in control agencies received an average of about 76 visits, whereas those in treatment agencies received only 58 visits--a difference of roughly 18 visits, or about 24 percent. Reductions in the number of visits by skilled nurses and home health aides explain more than 90 percent of this decrease; however, other services also experienced large and significant decreases. Skilled nursing visits fell from about 35 to 26 (a reduction of 26 percent), and home health aide visits fell from 30.5 to 23.2 (a reduction of 24 percent). Similarly, therapy services fell an average of 1.4 visits (19 percent), and medical social worker services fell an average of 0.4 visits (32 percent).

Within the at-risk period, prospective payment reduced the total visits per patient from 44.8 to 37.3, a decrease of 7.5 visits (17 percent), or about two-fifths of the decrease in the first year after admission. Skilled nursing services accounted for more than half this amount, falling an average of four visits per patient, or 19 percent. Home health aide services fell by close to three visits per

patient (17 percent), and medical social worker services fell by 0.3 visits (38 percent). The percentage of patients receiving one of these services dropped either slightly (in the case of skilled nursing) or not at all, suggesting that the reductions in service use stemmed from fewer visits to those who received visits. With respect to therapy services, the likelihood of an occupational therapy visit decreased significantly during the at-risk period, possibly because of a decline in its use as a supplement to physical therapy. However, the average number of therapy visits (by type and total) did not change significantly, which may have been due to more intensive use of therapy near the start of the care episode that offset later cutbacks.

Service provision continued to fall significantly after the initial at-risk period had ended. From the 5th through the 12th month after admission, the average number of visits fell by about one-third, from 30 to 20, an even greater decrease than had occurred during the at-risk period. Visits by skilled nurses and home health aides accounted for most of this drop, which is not surprising since these visits reflect the vast majority of services provided to longer-term patients. In contrast to the at-risk period, visits by physical therapists and occupational therapists fell significantly, from a combined average of about 2.2 visits to 1.4 visits. Medical social worker visits, which were provided only rarely after the at-risk period, also fell by a small but significant margin (about 0.1 visits per patient).

Earlier discharge is a key reason for all service reductions, particularly those occurring after the at-risk period had ended. Based on episodes observed over a maximum period of 480 days, the average episode length fell from 131 days to 98 days--a drop of more than one month. The percentage of patients with longer episodes (of various lengths) fell, while the proportion with shorter episodes increased, suggesting that treatment agencies reduced episodes of all types. In total, more than three-quarters of episodes in treatment agencies ended during the at-risk period, and a majority ended within only 60 days.

In addition to discharging patients sooner, treatment agencies also reduced the frequency of visits provided during both the at-risk period and outlier period. In each month after admission, patients with an episode of care continuing through the end of the month received an average of one to two visits less from treatment agencies than from control agencies. This difference is all the more striking when one considers that the mix of patients continuing care through a given month in treatment agencies was probably less healthy than those in control agencies, because the treatment agencies presumably shortened episodes by more quickly discharging the healthier patients (of all patients still receiving care at any point after admission). The similarity of this difference over time suggests that agencies were generally not distinguishing between the outlier and at-risk periods but, rather, pursued initiatives to reduce the number of visits at all points after admission.

Despite the tremendous response of treatment agencies to the demonstration incentives, we find no evidence that they tried to "game" the payment system in order to obtain additional profit. Given the dramatic reduction in episode lengths, for example, we might expect that at least some treatment agencies discharged patients earlier in order to readmit them into a second care episode. However, the rate of second episodes was fairly small and quite similar for agencies in both groups (roughly 10 percent by 16 months postadmission). We also find a very similar rate of episodes with only one or two visits (about four percent), contradicting the hypothesis that treatment agencies would admit a higher proportion of patients with little or no true care needs. Finally, we find no evidence that treatment agencies increased visits for longer-stay patients in order to obtain additional revenue from visits provided during the outlier period. In fact, the findings suggest that treatment agencies behaved during the outlier period precisely as they did during the at-risk period--by making significant and broad-based reductions in average visits and service length.

Although all impacts among various agency subgroups were large and significant (thus underscoring the main findings), we did find a number of important subgroup differences. One clear difference is that agencies that had provided a relatively large number of visits per episode before the demonstration began made larger reductions in the number of visits per patient than did relatively "low-use" agencies (about 24 visits per patient in the year after admission for high-use agencies, compared with 12 visits for low-use agencies). However, these decreases were actually quite similar in percentage terms (about 23 percent and 26 percent, respectively), because high-use (control) agencies provided many more visits than lower agencies in the absence of the demonstration. This difference was largely centered in home health aide visits and was almost certainly linked to the greater opportunity for high-use agencies to make large-scale reductions in services.

Most of the other agency subgroup differences appear to be associated with practice patterns. For example, for-profit agencies made larger cuts in visits than did non-profit ones, but this difference was due entirely to a higher proportion of for-profit agencies with high-use practice patterns prior to the demonstration. Thus, if a nonprofit and for-profit agency with the same practice patterns were placed on prospective payment, we would expect them to make similar cutbacks in services. Small agencies (those providing fewer than 30,000 visits before the demonstration began) and freestanding agencies also made greater reductions than did their counterparts; again, however, the difference in behavior appears to be explained by the greater likelihood that these agencies had high-use practice patterns prior to the demonstration.

A final difference, which is not related to practice patterns, is a far greater reduction in services by agencies above the cost limits compared with those below the limits. This difference may be explained by the particularly strong incentive that treatment agencies above the limits face to cut

visits, and the completely contrasting incentive that control agencies above the limits face to raise visits.

With respect to patient subgroups, we find that patients who had characteristics normally associated with a high level of service use had their average visits reduced by about twice the amount of those with relatively low expected service use. However, this difference in impacts (25 visits compared to 13 visits) was only marginally significant, and it was quite similar when measured as a proportion. Somewhat surprisingly, we find no significant differences between those who can take oral medications independently and those who cannot, or between those who have other caregivers available and those who do not. Patients who cannot take oral medications independently might be expected to have relatively more severe conditions; therefore, the similar service reductions experienced by these patients suggests that we carefully examine their outcomes in future reports.

Finally, the available data from the second demonstration year suggest that the impacts of prospective payment remained stable over time. However, this stability of impacts does not mean that agencies' behavior remained stable. To the contrary, both treatment and control agencies made large but offsetting reductions in visits and service length during the second year (compared with the first), resulting in no change in demonstration impacts. Whether these changes would eventually weaken or strengthen the demonstration impact is unclear. However, even if factors that reduced control agencies' service provision in year two (for example, managed care growth and investigations of fraud and abuse) continued to intensify, they would almost certainly not lead to the dramatic service reductions that treatment agencies achieved.

These findings proved to be highly robust. Impacts remained large and significant regardless of either the statistical model used or the manner in which we weighted the samples. Moreover, we find no evidence that our results are the result of incomplete service use data provided by treatment

group agencies, a few agencies with exceptionally high or low visits during the period, or unobserved differences between agencies or their patients.

B. LIMITATIONS OF THE ANALYSIS

Despite the significance and robustness of our findings, this study does have some limitations. Perhaps the most important potential limitation is the extent to which we can generalize our findings to home health agencies nationwide. Like any study design based on voluntary participation, the agencies in this demonstration may reflect the group most able (or most willing) to respond to the incentives of the intervention. If true, our demonstration impacts may overstate the actual reductions in service use that would take place nationally under a similar payment system. In addition, even if demonstration agencies had no particular predisposition to reduce the number of visits relative to similar agencies that are not participating in the demonstration, we may still overstate the impacts of prospective payment nationally because of other key differences between agencies in and out of the demonstration. For example, if agencies nationwide had relatively low-use practice patterns compared with demonstration agencies, we would probably overstate the effects of a nationwide program. Of course, if this difference was reversed, our impacts might actually understate these effects. Similarly, if control agencies were more responsive than agencies nationally to other pressures to contain or reduce their visits per episode, then our results might underestimate the national effects of a prospective payment system.

A key factor that suggests our results can be widely generalized is that agencies of *all types* made substantial reductions in their provision of services. Agencies of differing sizes, auspice, or practice patterns might be expected to differ both in their capacity and desire to respond to demonstration incentives; however, we find that each agency type made cuts of at least 18 percent

in their visits per patient. Thus, even if the mix of demonstration agencies differed significantly from agencies nationwide, we expect the impacts of a nationwide program to be substantial.

A second limitation, related to the issue of generalizability, is that any national program of prospective payment will differ from the one implemented for the demonstration. For example, agencies will not be protected from incurring financial losses under the national system being implemented, which could lead some agencies to respond more cautiously to program incentives. In addition, an agency's payment will not be based on its prior cost per episode. For many agencies, such as those with high-use practice patterns, this could actually lead to larger cuts than were observed under the demonstration, because they might have to reduce their costs per episode substantially to remain solvent.

A third limitation of this report is that it does not provide information on the consequences of reducing services. We address this issue in a series of other reports, by examining the effects of the demonstration on the quality of care and patient outcomes, patient selection and retention, and the utilization of other Medicare services. In addition, we examine the impacts of the demonstration on agencies' costs per visit and costs per episode, as well as on their profits and losses. A final report will combine the findings across outcome measures.

One final concern is that we examined impacts over only a short period. Indeed, because service provision by both treatment and control agencies fell significantly from the first to the second year of the three-year demonstration, we may not have an accurate measure of how prospective payment affects participating agencies over a longer period. For the most part, however, the findings from the first demonstration year provide all the information necessary to policymakers: within the first year of the demonstration, nearly all participating agencies in the treatment group made dramatic cuts in service use that may be directly attributed to the new payment system. A key

hypothesis that the demonstration was designed to investigate--whether prospective payment could curtail the tremendous growth in home health services--was thus strongly affirmed, despite the limited follow-up period.

C. POLICY IMPLICATIONS

The findings from this report provide key information on the effects of prospective payment at the very time when HCFA, under mandate from the 1997 Balanced Budget Act, must implement a nationwide prospective payment system. Two basic results from this study are perhaps most relevant to HCFA's efforts. First, prospective payment does appear to generate sizable reductions in service use, which may, in turn, lead to substantial savings to Medicare. Second, the use of a 120-day at-risk period followed by a prospectively set per-visit rate system (that is, the specific version of prospective payment evaluated for the demonstration) achieved precisely the desired effect, suggesting that it provides a sound starting point for implementing a national system.

1. Implications of the Reduction in Service Use

The large service reductions by treatment agencies suggest that a nationwide prospective payment system may lead to substantial savings relative to a system of cost-based reimbursement. However, such savings may be somewhat smaller proportionally than the overall reduction in visits (roughly 24 percent), because agencies' cost per visit may increase under prospective payment (Cheh and Trenholm 1998). Nonetheless, even if savings were only a fraction of this size, HCFA may have the opportunity to reduce spending on home health care while allowing agencies to prosper. For example, it could establish the prospective rates for an at-risk period as a proportion (say, 90 percent) of the historic average cost per episode under cost-based reimbursement, much as is done for payment to health maintenance organizations participating in the Medicare risk program. Our

forthcoming report on program effects on cost per visit will provide additional evidence on reasonable values for this proportion. After prospective payment has been implemented, the rates could be adjusted annually by the projected rate of increase in labor costs in the home health market. Additional changes to prospective rates could also be considered, depending on the profitability of agencies.

While similar cost savings might be achieved through direct regulation of the home health industry, prospective payment is likely to be a better approach. Consider, for example, the differences between a system of prospective payment and the recently enacted interim payment system (IPS).[1] Although IPS has led to sharp reductions in Medicare home health expenditures, these reductions have been achieved through caps imposed on per-beneficiary spending and per-visit costs. Effectively, these caps limit not only the level of care provision but also the mode of savings as some approaches to cost-cutting may not benefit the agency. (For example, agencies would not want to generate savings through approaches that might drive its per-visit costs above the cap). In contrast, a per-episode prospective payment system offers great flexibility both in how agencies serve their patients and how they generate savings. Moreover, because agencies' profits rise as their costs decline, prospective payment creates the incentive to generate as much savings as possible (short of degrading care quality). Prospective payment thus provides an opportunity for savings that is similar to what might be imposed under a more regulated system of payment, while allowing agencies greater flexibility to provide care that is tailored to the needs of individual patients.

---

[1] The 1997 Balanced Budget Act mandated an interim payment system (IPS) until the date that HCFA has implemented prospective payment. Under IPS, agencies received the lowest of (1) their actual allowable costs; (2) their aggregate cost per visit, lowered to 105 percent of the established Section 223 Limits for the region; and (3) a new aggregate per-beneficiary limit.

2. Implications for the Design of a National Program

Although a voluntary demonstration cannot address all the issues that must be considered in the design of a national payment system, the basic demonstration design--a 120-day at-risk period followed by prospective per-visit payments--appears to be a promising approach. Indeed, based on the complete set of study findings to date, the demonstration design may be capable of generating large savings relative to a system of cost-based reimbursement while maintaining both access to care and care quality. The demonstration design thus provides a sound starting point for implementing a national payment system, which can be modified as necessary to address concerns that were beyond the scope of the demonstration.

One modification that HCFA may wish to consider, for example, is a "blended" payment system that includes both an episodic component and a per-visit (cost-based) component for each at-risk period. Although agencies would presumably have their payments adjusted on the basis of their case mixes, this payment system may help ensure that agencies are sufficiently compensated for higher-need patients. The blended payment could be implemented in several ways. For example, agencies could have their episodic payments adjusted retrospectively according to the number and types of visits provided to a given patient. The key benefit of this approach is that agencies would be more accurately reimbursed for their patient costs (because the payment accounts for actual difference in utilization) but would still have an incentive to control utilization, because the added revenue from additional visits would be based on only a fraction of the agency's per-visit costs. However, an important downside of this approach is the operational burden imposed by the existence of two different payment systems.

Per-episode payment creates several potential incentives for gaming, suggesting that it might be important to establish formal systems to monitor agencies' behavior. A formal system should be

able to identify agencies that have an inordinate number of short episodes, provide a disproportionate share of outlier visits, or readmit a particularly high rate of patients into subsequent episodes of care.

Finally, because the long-run implications of prospective payment are difficult to anticipate, it will be important to continually monitor the evolution of the home health industry and its effect on patient access and outcomes. In particular, given a reduced payment, some home health agencies may face consolidation or closure, leading to the possible loss of access for some home health patients. To the extent that HCFA can anticipate these changes, it can adopt policies to ensure that patients continue to receive high-quality home health care.

D. SUMMARY

Prospective payment has been found to be a highly effective tool for controlling service provision by home health agencies. Although controlling service provision does not necessarily equate to cost savings for HCFA, in all likelihood, the government should obtain substantial savings from a prospective payment system relative to cost-based reimbursement. If prospective payment does not harm patients' access or outcomes or substantially raise other Medicare costs, implementing a prospective payment system would appear to be a sound approach to controlling home health costs.

REFERENCES

Brown, Randall, Barbara Phillips, Christine Bishop, Amy Klein, Grant Ritter, Craig Thornton, Peter Schochet, and Kathleen Skwara. "The Effects of Predetermined Payment Rates for Home Health Care." Princeton, NJ: Mathematica Policy Research, Inc., December 1995.

Cheh, Valerie, and Christopher Trenholm. "Preliminary Report: The Impact of Prospective Payment on The Cost Per Episode: Striking The Balance Between Decreasing use and Increasing Cost (Draft)." Princeton, NJ: Mathematica Policy Research, Inc., February 1998.

Chen, Arnold, and Helaine Noveck. "Preliminary Report: The Impact of Prospective Payment on Medicare Home Health Quality of Care. Princeton, NJ: Mathematica Policy Research, Inc., January 1998.

Health Care Financing Review. Statistical Supplement, 1998.

Health Care Financing Review. Medicare and Medicaid Statistical Supplement, 1996.

Phillips, Barbara, and Rachel Thompson. "Transition Within a Turbulent System: An Analysis of the Initial Implementation of the Per-Episode Home Health Prospective Payment Demonstration." Draft report. Princeton, NJ: Mathematica Policy Research, Inc., February 1997.

Phillips, Barbara R., Randall Brown, Jennifer Schore, Amy C. Klein, Peter Z. Schochet, Jerrold W. Hill, and Craig Thornton. "Evaluation of the Per-Episode Home Health Prospective Payment Demonstration Design." Report submitted to the Health Care Financing Administration. Princeton, NJ: Mathematica Policy Research, Inc., September 1995.

Phillips, Barbara R., Randall S. Brown, Jennifer L. Schore, Amy C. Klein, Peter Z. Schochet, Jerrold W. Hill, and Dexter O. Chu. "Case-Mix Analysis Using Demonstration Data Home Health Prospective Payment Demonstration." Draft report. Princeton, NJ: Mathematica Policy Research, Inc., December 1992.

PriceWaterhouseCoopers LLP. "Congress Changes Medicare Reimbursement for Home Health Services: What Does It Mean for Home Health Providers." *Health Policy Series,* No. 6, September 1998.

Teplitsky, Sanford V., and Mary Ann Janson (eds.). *Home Health & Hospice Manual: Regulations and Guidelines*. Owings Mills, MD: National Health Publishing, 1985-1992.

Trenholm, Christopher A. "Effects of Per-Episode Prospective Payment for Medicare Home Health Care on Patient Selection and Retention." Draft report. Princeton, NJ: Mathematica Policy Research, Inc., June 8, 1999.

U.S. General Accounting Office. *Medicare Home Health Care. Prospective Payment System Will Need Refinement as Data Becomes Available.* Washington, DC: GAO, April 2000.

U.S. General Accounting Office. *Medicare Health Benefit: Impact of Interim Payment System and Agency Closures on Access to Services.* Washington, DC: GAO, September 1998.

APPENDIX A

SUMMARY OF CASE-MIX ADJUSTMENT FOR PAYMENTS DURING THE DEMONSTRATION

During the demonstration, agency payment rates are adjusted for severity using the Home Health Utilization Groups (HHUGs) Classification system. This system, which was specifically developed for the demonstration, classifies patients into 1 of 18 mutually exclusive cells on the basis of the following information:

- Whether the patient has an intervening hospital stay in an acute-care hospital during the 120 days following a home health admission
- Whether the patient was impaired in at least four of five activities of daily living (bathing, eating/tube feeding, dressing, toileting/elimination, and transferring) at home health admission
- Whether wound care was planned for the patient at home health admission
- Whether the patient was discharged from a hospital within the 14 days preceding home health admission
- Whether the patient had stage 3 or stage 4 decubitus ulcer at home health admission
- Whether the patient had cancer at home health admission that affected current treatment or personal-care needs
- Whether the patient has had a cerebrovascular accident that affected current treatment or personal-care needs
- Whether the patient had diabetes that affected current treatment or personal-care needs

As with payment rates, demonstration case-mix adjustment is based on the agency's base-period experience. Information on the characteristics listed above was collected on each patient admitted to each agency in the 90 days ("base quarter") before it began to operate under the demonstration. The same information was collected throughout the demonstration in the "remarks" section of the Uniform Bill, HCFA form 92 (UB-92).

Using the base quarter information, Abt calculated a base quarter index for each agency as follows. First, Abt calculated a category weight for each of the 18 case-mix groups by dividing the

average cost in each group by the agency's overall average cost.[1] Abt then multiplied each category weight by the percentage of episodes that fell into each case-mix category during the base quarter, and summed across all 18 categories. The result is a base quarter index for each agency, which is always equal to 1.

To obtain the annual case-mix-adjusted rates, Abt calculated a similar index for the demonstration years by multiplying the category weight from the base year times the percentage of episodes that fell into each case-mix category during the relevant year, and then summing across all 18 categories.

The demonstration year index for each agency is then divided by the agency's base quarter index and multiplied by the agency's base year episode rate. This results in the case-mix-adjusted episode rate.

---

[1] If an agency did not have any observations in a particular category during the base quarter, the average cost for that casemix group calculated across all demonstration agencies was used instead of an agency-specific cost.

APPENDIX B

DATA QUALITY

As first noted in Chapter II, we identified nine agencies that may not have submitted a large portion of their bills for services rendered during the demonstration. Because this non-billing leads the observed number of visits to be smaller than the actual number, it could cause serious misstatements about the effects of prospective payment on agencies' service provision. We therefore exclude all nine of these agencies from our analysis sample.

In this appendix, we summarize the process that led to the exclusion of these nine agencies, and we then conduct sensitivity analyses to determine how this and any other non-billing by demonstration agencies may have influenced our main findings. In brief, these sensitivity analyses indicate no evidence that the large impacts of prospective payment presented in this report are due to the necessary exclusion of agencies or to any additional missing bill data.

A. IDENTIFYING AGENCIES WITH MISSING OR ERRONEOUS CLAIMS

We first identified a potential problem with billing by some demonstration agencies through an initial review of the data for the first year of the demonstration conducted in the late summer of 1996. During this review, we discovered two agencies (both in the treatment group) that displayed extremely large reductions in average visits per patient from the predemonstration to the demonstration, as well as very high rates of one or two visit episodes. After contacting these agencies, we learned that they had indeed not submitted a large fraction of their bills for the first demonstration year. One of these two agencies clearly indicated that its behavior was a direct result of receiving prospective payment. Specifically, because it was paid per episode and upon acceptance of the initial bill, the agency believed that incurring the costs of billing for visits rendered later in the 120-day at-risk period was not in its best interest.[1] The other agency indicated that they had

---

[1] Because of the complicated nature of the cost-reporting process and the calculation of losses, (continued...)

fallen seriously behind in their billing because of administrative problems; however, they further indicated that they were particularly slow in submitting interim bills for which they would not be paid. Thus, the demonstration payment method appeared to be affecting the behavior of this agency as well.

The discovery of non-billing by these two agencies raised immediate concerns that other agencies may also not be submitting bills, but less noticeably. To measure the extent and source of any billing problems among the demonstration agencies, we initially conducted two data analyses. The first examined bill records and checked for the frequency of missing interim bills. Using episodes that were at least 60 days long, we identified those episodes that included only an initial bill and a discharge bill, and we compared the proportion of these episodes across agencies. The second analysis used Medicare claims data and the quality assurance (QA) data. In this analysis, we compared the bill-through date of the last Medicare bill of the episode with the last day of service as recorded on the QA forms. If agencies failed to submit all their bills, then the last day of service from the QA form would be after the discharge date obtained from the bills.

Each of these analyses had limitations that hampered our ability to identify agencies with possible billing problems. In the case of searching for missing interim bills, we found that submission of only the initial and final bills was commonplace in many agencies and often involved more than 20 percent of their episodes, regardless of treatment status. This made it difficult to distinguish agencies that had truly failed to submit bills from those who had not. In the case of analyzing the QA data, we had information only on episodes starting after April 1996--several months after the last agencies had enrolled in the demonstration. As a result, these data included

---

[1] (...continued)
there is some financial advantage to bill completely. However, not all agencies understood the advantage at this point in the demonstration.

only a small fraction of many agencies' episodes, which led us to question whether they were representative of all episodes. In addition, early in the demonstration, not all agencies submitted QA data systematically, which led to an even smaller (and potentially more nonrepresentative) sample of episodes within each agency.

Given these limitations, we decided to contact all the demonstration agencies in the fall of 1997 and request information about their billing practices. From these interviews and the previous data analyses, we identified twenty-five agencies (in addition to the original two) that likely experienced some problems in billing during the first year of the demonstration. A total of nine of these agencies were in the control group, while the remaining sixteen were in the treatment group. In most cases, these problems did not appear to be substantial; however, five agencies in the treatment group did display evidence of substantial non-billing that, in each case, was linked to the method of payment. Namely, the agencies simply failed to submit (or resubmit) bills for services provided in the first 120 days because they had already been reimbursed through the episode payment.

At the second round of site visits conducted in the spring and summer of 1998, we followed up on our previous findings by asking each demonstration agency to describe any past or ongoing problems that they had experienced with billing, as well as if and when these problems had been rectified.[2] Of the seven treatment agencies previously identified with serious billing problems, we were not able to conclude satisfactorily that any had fixed their billing problems from the first year by the point that we downloaded the bill data to construct our analysis sample. In addition, we identified two other agencies (one treatment and one control) with substantial billing problems that may have dated back to the first year of the demonstration. We also identified three other treatment

---

[2] Any agencies not included in the second round of site visits were interviewed again by telephone.

agencies that had relatively minor billing delays, estimated at less than five percent of their total visits.

In summary, our investigation led us to identify nine agencies that may not have billed for a substantial share of services by the time we constructed our analysis sample for this report. Eight of these agencies were in the treatment group, and one was in the control group. In addition, we identified 23 agencies with more minor billing problems in the first demonstration year, 14 in the treatment group and 9 in the control group. The former group was excluded from our analysis sample, while the latter group was not excluded.

B. THE SENSITIVITY OF OUR RESULTS TO BILLING PROBLEMS

The nine agencies that displayed serious billing problems in the first demonstration year were dropped from the analysis because of concerns that their inclusion may seriously bias our impact estimates. Since all but one of these nine agencies is in the treatment group, including these agencies would probably have overstated demonstration impacts, since the mean for visits for treatment agencies would likely have fallen. This assumption should be tested, however, particularly since the agencies that were dropped reflected a non-random group. If we find that the impacts did not increase when including a large number of treatment agencies known to have missing bill data, it may suggest that our main findings are overstated by focusing on a (non-random) subset of agencies. Conversely, if the impacts increase by a great deal when including these agencies, it may suggest that our main impacts are actually understated. This latter inference would never be conclusive, however, since we cannot separate the effects of non-billing from the effects of actual reductions in service use.

The 23 agencies (14 in the treatment group and 9 in the control group) that displayed more minor problems with billing were included in the analysis because any missing bill data was not

believed to be significant enough to bias our impact estimates. This assumption should certainly be tested, however. If we find a significant change in the impacts when excluding these agencies, it may suggest that our main impact estimates are biased by including agencies that have a portion of missing bills. Conversely, if we observe little change in the main impact estimates, it strongly suggests that non-billing has not biased our results.

To investigate these various hypotheses, we reestimated selected impacts of prospective payment on three additional samples:

1. All 87 agencies for which we have sufficient demonstration data, regardless of whether they show evidence of unsubmitted bills.[3] Results from this regression provide a test of whether our main impact estimates may have been affected by the necessary exclusion of a non-random subset of agencies with incomplete billing.

2. The 56 agencies that displayed no evidence of billing problems during the first demonstration year. Results from this regression provide an indication of whether our main impact estimates have been affected by missing billing data, believed to be minor.

3. The 66 agencies that displayed no evidence of billing problems in the first demonstration year that could possibly be attributed to prospective payment. This sample excludes all treatment agencies that displayed any problems with billing, but it does not exclude any control agencies with billing problems. As such, it should provide the most conservative estimate of demonstration impacts.

The impact estimates from the first sensitivity sample (of 87 agencies) are somewhat larger than for the main analysis sample, which contradicts the hypothesis that the non-random exclusion of agencies has caused our main impacts to be overstated (Appendix Table B.1). In general, the difference in impacts is fairly modest, which suggests that our main impact estimates are also not materially understated by excluding these agencies. With respect to total visits, for example, the

---

[3] This sample includes only eight of the nine agencies dropped from the main analysis for likely non-billing, because one (treatment) agency had too few admissions (nine) in the first demonstration year to be investigated in regression models that weight agencies equally.

TABLE B.1

IMPACTS OF PROSPECTIVE PAYMENT WHEN NOT EXCLUDING AGENCIES WITH EVIDENCE OF MISSING BILL DATA

	Excluding No Agencies Because of Missing Bill Data			Main Analysis
Outcome	Treatment Group Mean[b]	Control Group Mean	Difference [b] (*p*-value)[c]	Impact[b] (*p*-value)[c]
Services by One Year from Admission[a]				
Total Visits	54.2	75.2	-21.0*** (0.00)	-18.3*** (0.00)
Skilled Nursing Visits	24.8	34.8	-10.0*** (0.00)	-9.2*** (0.00)
Home Health Aide Visits	21.3	30.1	-8.8*** (0.00)	-7.3*** (0.00)
Physical Therapy Visits	6.2	7.5	-1.3** (0.04)	-1.0 (0.12)
Length of Service (from Admission)[a]				
Days Until Last Visit (Maximum = 366)	107.5	133.6	-26.1*** (0.00)	-24.7*** (0.00)
Probability of Completing Services Within First 120 Days (Percentage)	70.8	61.6	9.2*** (0.00)	8.8*** (0.00)
Sample Size:				
Patients	**31,565**	**26,460**		**52,133**
Agencies	**46**	**41**		**79**

SOURCE: Medicare claims data.

NOTE: Individual observations (patient records) have been weighted to give agencies equal representation in the calculations.

[a] Based on a patient's first admission to a home health agency participating in the demonstration.

[b] Regression-adjusted (through ordinary least squares and logit models) to control for preexisting differences between treatment and control agencies.

[c] The *p*-value corresponds to a test of whether the treatment-control difference (impact) is statistically different from zero. It is based on standard errors inflated to account for the effects of clustering and weighting.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.
***Significantly different from zero at the .01 level, two-tailed test.

impact with the sensitivity sample was about three visits greater than with the main analysis sample, a small enough difference to be explained by (a large share of) missing bills among the excluded agencies.

The impact estimates from the second sensitivity sample (of 56 agencies) are very similar to the main impact estimates, which suggests strongly that our main findings have not been affected by any non-billing by agencies included in the analysis sample. (Appendix Table B.2) The impacts on total visits, for example, are actually slightly larger when excluding agencies with any evidence of missing bills, a somewhat surprising result since a majority of these agencies were in the treatment group.

The impact estimates from the third sensitivity sample (of 65 agencies) are also very similar to the main impact estimates, which provides further evidence that our reported findings are unaffected by any non-billing among agencies included in the analysis sample (Appendix Table B.3). In fact, all reported impacts are slightly larger with the sensitivity sample, a surprising result since it excludes only the treatment agencies identified for non-billing.

In summary, we conclude from these analyses that the existence of missing bill data by agencies is not responsible for the large reported impact of the demonstration payment method on the provision of services. When comparing our main impact estimates to those based on samples that either include or exclude agencies with possible non-billing in the first demonstration year, the differences in impacts are small and in the wrong direction to suggest that our main finding have overstated the reductions in service use brought about by prospective payment.

TABLE B.2

IMPACTS OF PROSPECTIVE PAYMENT WHEN EXCLUDING ALL AGENCIES WITH ANY EVIDENCE OF MISSING BILL DATA

	Excluding All Agencies with Evidence of Missing Bill Data			Main Analysis
Outcome	Treatment Group Mean[b]	Control Group Mean	Difference[b] (*p*-value)[c]	Impact[b] (*p*-value)[c]
Services by One Year from Admission[a]				
Total Visits	60.5	78.2	-17.7*** (0.00)	-18.3*** (0.00)
Skilled Nursing Visits	26.9	35.7	-8.8*** (0.00)	-9.2*** (0.00)
Home Health Aide Visits	26.0	32.5	-6.5** (0.02)	-7.3*** (0.00)
Physical Therapy Visits	5.6	7.1	-1.5 (0.11)	-1.0 (0.12)
Length of Service (from Admission)[a]				
Days Until Last Visit (Maximum = 366)	110.7	136.9	-26.2*** (0.00)	-24.7*** (0.00)
Probability of Completing Services Within First 120 Days (Percentage)	69.8	60.5	9.3*** (0.00)	8.8*** (0.00)
Sample Size:				
Patients	**16,427**	**18,556**		**52,133**
Agencies	**25**	**31**		**79**

SOURCE: Medicare claims data.

NOTE: Individual observations (patient records) have been weighted to give agencies equal representation in the calculations.

[a] Based on a patient's first admission to a home health agency participating in the demonstration.

[b] Regression-adjusted (through ordinary least squares and logit models) to control for preexisting differences between treatment and control agencies.

[c] The *p*-value corresponds to a test of whether the treatment-control difference (impact) is statistically different from zero. It is based on standard errors inflated to account for the effects of clustering and weighting.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.
***Significantly different from zero at the .01 level, two-tailed test.

TABLE B.3

IMPACTS OF PROSPECTIVE PAYMENT WHEN EXCLUDING ONLY TREATMENT AGENCIES WITH EVIDENCE OF MISSING BILL DATA

	Excluding Only Treatment Agencies with Evidence of Missing Bill Data			Main Analysis
Outcome	Treatment Group Mean[b]	Control Group Mean	Difference[b] (*p*-value)[c]	Impact[b] (*p*-value)[c]
Services by One Year from Admission[a]				
Total Visits	55.9	75.2	-19.3*** (0.00)	-18.3*** (0.00)
Skilled Nursing Visits	25.3	34.8	-9.5*** (0.00)	-9.2*** (0.00)
Home Health Aide Visits	22.5	30.1	-7.6*** (0.01)	-7.3*** (0.00)
Physical Therapy Visits	6.1	7.5	-1.4* (0.07)	-1.0 (0.12)
Length of Service (from Admission)[a]				
Days Until Last Visit (Maximum = 366)	106.3	133.6	-27.3*** (0.00)	-24.7*** (0.00)
Probability of Completing Services within First 120 Days (Percentage)	71.3	61.6	9.7*** (0.00)	8.8*** (0.00)
Sample Size:				
Patients	**16,415**	**26,460**		**52,133**
Agencies	**25**	**41**		**79**

SOURCE: Medicare claims data.

NOTE: Individual observations (patient records) have been weighted to give agencies equal representation in the calculations.

[a] Based on a patient's first admission to a home health agency participating in the demonstration.

[b] Regression-adjusted (through ordinary least squares and logit models) to control for preexisting differences between treatment and control agencies.

[c] The *p*-value corresponds to a test of whether the treatment-control difference (impact) is statistically different from zero. It is based on standard errors inflated to account for the effects of clustering and weighting.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.
***Significantly different from zero at the .01 level, two-tailed test.

APPENDIX C

COEFFICIENT ESTIMATES FROM THE REGRESSION ON TOTAL VISITS PER PATIENT YEAR

The complete list of coefficients from the regression on visits per patient year (shown in Appendix Table C.1) indicates that a number of patient- and agency-level characteristics have significant effects.

Nearly all the patient characteristics associated with case mix (including ADLs, medical conditions, and length of prior hospitalization) have significant effects in the expected direction. Among these case-mix variables, those having the largest impact are diabetes (15.6 visits) and stroke (15.0 visits), as well as wound care (14.4 visits) and bathing (13.4 visits). Most of the demographic variables (unrelated to case mix) also have a significant effect on visits. All else equal, older age groups (75 and over) had roughly 5 to 10 more visits per patient year than those 65 to 74, while those with age as original reason for entitlement had 6.6 fewer visits than those whose original reason for entitlement is disability. Finally, whites had 9.4 fewer visits than nonwhites.

Aside from treatment status, only a few of the agency-level variables had significant impacts on visits. By far the largest of these is the agency practice pattern.[1] For example, given a one standard deviation increase in the agency practice pattern (about 0.31), the number of visits per patient year increased by about 11.1 visits. The state in which an agency is located also had a significant effect on visits. Controlling for other agency characteristics, agencies in Massachusetts and Florida provide far more visits per episode (25.6 and 14.5, respectively) than those in California (the base category); however, there was no significant difference in visits between agencies in California and those in Texas or Illinois. Somewhat surprisingly, there was no significant difference in the provision of visits depending on whether an agency is nonprofit or proprietary, small or large, or freestanding or hospital-based.

---

[1] As described in Chapter II, this variable reflects the predemonstration, case-mix-adjusted ratio of the average visits per episode in a given agency to the average visits per episode in all agencies.

TABLE C.1

COEFFICIENT ESTIMATES FROM ORDINARY LEAST SQUARES REGRESSION ON TOTAL VISITS IN THE YEAR AFTER ADMISSION[a]

	Impacts on the Number of Visits per Patient Year	(*p*-value)[b]
Agency Received Prospective Payment	-18.3***	(0.00)
Age		
Younger than 65	-0.2	(0.98)
75 to 84	4.9**	(0.03)
85 or older	9.7**	(0.02)
Female	0.6	(0.78)
White	-9.4**	(0.02)
Original Reason for Medicare: Old Age	-6.6	(0.11)
Has Medicare Buy-In	0.2	(0.92)
Medical Conditions		
Cancer	-9.1***	(0.01)
Diabetes	15.6***	(0.00)
Cerebrovascular accident (stroke)	15.0***	(0.00)
Decubiti stage 3 or 4	11.7**	(0.03)
Need for Complicated Wound Care[c]	14.4***	(0.00)
Functional Limitations[d]		
Bathing	13.4***	(0.00)
Eating	7.0**	(0.02)
Dressing	4.4***	(0.01)
Toileting	11.2***	(0.00)
Transferring	6.8***	(0.01)
Preadmission Location: Hospital	-13.8***	(0.00)
Had Medicare for Less than Six Months	0.2	(0.97)
Practice Patterns[e]	35.9***	(0.00)
Hospital-Based Agency	-1.1	(0.78)
For-Profit Agency	5.6	(0.13)
Chain Member	-3.5	(0.45)
Agency Provided Fewer than 30,000 Visits in Base Year	2.9	(0.54)
Agency Located in Urban Area	-6.4	(0.42)
Agency State		
Florida	14.5***	(0.00)
Illinois	-3.4	(0.59)
Massachusetts	25.6***	(0.00)
Texas	6.1	(0.27)

TABLE C.1 *(continued)*

	Impacts on the Number of Visits per Patient Year	(*p*-value)[b]
County-Level Means		
Number of nursing home beds per 100 persons over age 65	4.1***	(0.00)
Number of physicians per 10,000 persons	-0.2	(0.49)
Hospital occupancy rate	10.9	(0.65)
Length of Hospital Stay During Two Weeks Before Home Health (Days)[f]	0.3	(0.33)
Any Skilled Nursing Facility Stay During Two Weeks Before Home Health	3.9	(0.15)
Total Medicare Part A Reimbursement During Six Months Before Home Health (in Thousands of Dollars)[g]	0.3**	(0.02)
Total Number of Home Health Visits During Six Months Before Demonstration Admission	0.5***	(0.00)
In an HMO at Any Time During the Six Months Prior to Admission	9.8	(0.27)
Had Medicare as a Second Payer at Any Time During the Six Months Prior to Admission	23.6**	(0.04)
Admission Between 5/1/95 and 9/30/95	12.5***	(0.01)
Admission Between 10/1/95 and 12/31/95	3.7	(0.27)
Admission Between 4/1/96 and 6/30/96	0.2	(0.96)
Admission Between 7/1/96 and 9/30/96	-3.9	(0.22)
Admission Between 10/1/96 and 12/31/96	-4.6	(0.34)
Intercept	-14.6	(0.39)
Sample Size	**52,133**	--

SOURCE: Control variables constructed from Medicare claims data, the enrollment database (EDB), and the case-mix data for the demonstration. See Chapter II for further information.

NOTE: Individual observations (patient records) have been weighted to give agencies equal representation in the calculations.

[a] Reflects a patient's first admission to a home health agency participating in the demonstration.

[b] The *p*-value corresponds to a test of whether the impact shown is statistically significant from zero. It is based on standard errors inflated to account for the effects of clustering and weighting.

[c] Patient has wound that requires soaking, irrigation, or debridement.

[d] Patient requires some human assistance with or does not participate in activity.

[e] A case-mix-adjusted index of the average number of visits received by an agency's patients relative to the average number provided by other demonstration agencies.

[f] If patient was not hospitalized within the two weeks before home health, days are set to zero.

[g] Includes reimbursement for inpatient, skilled nursing facility, hospice, and nondemonstration home health paid under Medicare Parts A and B.

*Significantly different from zero at the .10 level, two-tailed test.
**Significantly different from zero at the .05 level, two-tailed test.
***Significantly different from zero at the .01 level, two-tailed test.

CMS LIBRARY
3 8095 00012877 3